

Noether's Theorem Under the Legendre Transform

by

Jonathan Herman

A research paper
presented to the University of Waterloo
in fulfilment of the
research paper requirement for the degree of
Master of Mathematics
in
Pure Mathematics

Waterloo, Ontario, Canada, 2014

© Jonathan Herman, 2014

Abstract

In this paper we demonstrate how the Legendre transform connects the statements of Noether's theorem in Hamiltonian and Lagrangian mechanics. We give precise definitions of symmetries and conserved quantities in both the Hamiltonian and Lagrangian frameworks and discuss why these notions in the Hamiltonian framework are somewhat less rigid. We explore conditions which, when put on these definitions, allow the Legendre transform to set up a one-to-one correspondence between them. We also discuss how to preserve this correspondence when the definitions of symmetries and conserved quantities are less restrictive.

Contents

1	Introduction	1
2	Symplectic Geometry	3
2.1	Symplectic Vector Spaces	3
2.2	Symplectic Manifolds	4
2.3	When is a Diffeomorphism a Symplectomorphism?	6
2.4	Canonical Symplectic Structure of Cotangent Bundles	7
2.5	Lifting a Diffeomorphism	9
2.6	Constructing Symplectomorphisms	13
2.7	Applications to Geodesic Flow	15
3	Lagrangian Mechanics	20
3.1	Conservative and Central Forces	20
3.2	The Calculus of Variations and the Euler-Lagrange Equations	24
3.3	Examples of Lagrangian Systems	26
4	Hamiltonian Mechanics	31
4.1	Hamiltonian Vector Fields	31
4.2	Hamilton's Equations	32
4.3	The Poisson Bracket	34
5	The Legendre Transform	36
5.1	The Legendre Transform on a Vector Space	36
5.2	The Legendre Transform on Manifolds	41
5.3	The Legendre Transform Relates Lagrangian and Hamiltonian Mechanics	41
6	Noether's Theorem	49
6.1	Noether's Theorem in Lagrangian Mechanics	49
6.2	Noether's Theorem in Hamiltonian Mechanics	53
6.3	Noether's Theorem Under the Legendre Transform	55
6.4	The Converse of Noether's Theorem in the Lagrangian Setting	58
6.5	Relaxing the Definitions of Symmetries and Conserved Quantities	61

1 Introduction

This paper studies the theorem that Emmy Noether published in 1918, which provides a mathematical way to see connections between ‘symmetries’ and ‘conserved quantities’. As we shall see, Noether’s theorem can be stated in both the Lagrangian and Hamiltonian frameworks. In Section 6 we demonstrate how the Legendre transform relates these statements and furthermore how, under specific requirements, it gives a one-to-one correspondence between the respective notions of symmetry and conserved quantity.

Section 2 is dedicated to introducing the tools needed from symplectic geometry to formulate Hamiltonian mechanics. In Section 1.6 we will see how geodesic flow on a Riemannian manifold arises as a symplectomorphism generated by a specific diffeomorphism. In Section 5.3 we apply the Legendre transform to this setup and recover an equivalent way to define geodesic flow in the Hamiltonian framework.

Section 3 gives an introduction to Lagrangian mechanics. In particular, we derive the Euler-Lagrange equations using tools from the calculus of variations. A Lagrangian is just a smooth function on the tangent bundle and we will see that when this function is a ‘natural Lagrangian,’ the Euler-Lagrange equations are equivalent to Newton’s second law. We also demonstrate how the Euler-Lagrange equations are a generalization of Newton’s second law; in particular, the Euler-Lagrange equations hold in non-inertial reference frames. We give many examples of Lagrangian systems and then translate these systems to the Hamiltonian framework in Section 5.

In Section 4 we use the tools introduced in Section 2 to study some basic notions in Hamiltonian mechanics. As mentioned above, the main object of study in Lagrangian systems is the Lagrangian, which is just a smooth function on the tangent bundle TM . In Hamiltonian mechanics the main object is the Hamiltonian, which is just a smooth function on the cotangent bundle T^*M . In Section 2.4 we show how the cotangent bundle always has a canonical symplectic structure and so we see that an advantage of Hamiltonian mechanics is that it incorporates the use of tools from symplectic geometry.

After introducing Lagrangian and Hamiltonian mechanics, Section 5 demonstrates how the two formulations are equivalent under the Legendre transform. Given a Lagrangian $L \in C^\infty(TM)$ we get an induced map called the Legendre transform, which we denote by Φ_L , from TM to T^*M . Similarly, given a Hamiltonian $H \in C^\infty(T^*M)$ we get the induced Legendre transform $\Phi_H : T^*M \rightarrow TM$. Under certain conditions, which we discuss, the Legendre transform is a diffeomorphism. We use the Legendre transform to translate examples given in Section 3 and Sections 4 into the opposing frameworks. In particular, we will see how the Legendre transform takes motions in one framework to motions in the other.

In Section 6 we study Noether’s theorem in both the Lagrangian and Hamiltonian frameworks. We give physical examples in both settings to demonstrate the power of this theorem. We then show

how the statements of Noether's theorem can be translated, under the Legendre transform, from one framework to the other. We give examples of how the Legendre transform takes symmetries to symmetries and conserved quantities to conserved quantities. With the definitions given, we use the Laplace-Runge-Lenz vector to show how the notions of symmetry and conserved quantity are not in one-to-one correspondence. However, we fix this problem by putting restrictions on the symmetries and conserved quantities. Lastly, we discuss the problem of how to make the correspondence one-to-one when the definitions are more general.

Throughout this paper we will use the Einstein summation convention.

2 Symplectic Geometry

We discuss here the concepts in symplectic geometry which will be needed to formulate Hamiltonian mechanics in Section 4. In Riemannian geometry, manifolds are equipped with a non-degenerate symmetric quadratic form, whereas in symplectic geometry the non-degenerate quadratic form is required to be skew-symmetric. Although there are some similarities between symplectic and Riemannian geometry, as we shall see there are also some vast differences.

2.1 Symplectic Vector Spaces

Let V be an m dimensional real vector space and $\Omega : V \times V \rightarrow \mathbb{R}$ a skew-symmetric bilinear map. Let $U = \{u \in V ; \Omega(u, v) = 0 \text{ for all } v \in V\}$. Suppose that $\dim U = k$ and that $\{u_1, \dots, u_k\}$ is a basis. Recall the standard form theorem for skew-symmetric bilinear maps:

Theorem 2.1. (Standard Form for Skew-Symmetric Bilinear Maps) With U, V and Ω as above, we can find $n \in \mathbb{N}$ and a basis $u_1, \dots, u_k, e_1, \dots, e_n, f_1, \dots, f_n$ of V such that

$$\begin{aligned} \Omega(u_i, v) &= 0 && \text{for all } i \text{ and for all } v \in V \\ \Omega(e_i, e_j) &= 0 = \Omega(f_i, f_j) && \text{for all } i, j \\ \Omega(e_i, f_j) &= \delta_i^j && \text{for all } i, j \end{aligned}$$

Proof. This is a fairly straightforward induction proof. See [1], page 3 for details. \square

It follows that, with respect to this basis, the matrix representation of Ω is

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \text{Id} \\ 0 & -\text{Id} & 0 \end{bmatrix}$$

Definition 2.2. The bilinear map Ω is said to be **symplectic (or non-degenerate)** if $U = \{0\}$. If this is the case then the pair (V, Ω) is called a **symplectic vector space** and $\{e_1, \dots, e_n, f_1, \dots, f_n\}$ is called the corresponding **symplectic basis**.

It follows from the theorem that any symplectic vector space is necessarily even dimensional and the corresponding skew-symmetric bilinear map is of the form

$$\begin{bmatrix} 0 & \text{Id} \\ -\text{Id} & 0 \end{bmatrix}$$

Example 2.3. (Symplectic Vector Space Prototype)

The simplest example of a symplectic vector space is $(\mathbb{R}^{2n}, \Omega_0)$ where Ω_0 is defined such that $e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1, 0, \dots, 0)$ together with $f_1 = (0, \dots, 0, 1, 0, \dots, 0), \dots, f_n = (0, \dots, 0, 1)$ form a symplectic basis. The reason this symplectic vector space is referred to as a prototype is given by the Darboux theorem, which is stated in section 2.2.

Definition 2.4. Let (V_1, Ω_1) and (V_2, Ω_2) be symplectic vector spaces. A linear isomorphism $\varphi : V_1 \rightarrow V_2$ is called a **symplectomorphism** if $\varphi^* \Omega_2 = \Omega_1$. Here φ^* is the pullback of φ meaning that $(\varphi^* \Omega_2)(u, v) = \Omega_2(\varphi(u), \varphi(v))$.

In the same way that a Riemannian metric induces the musical isomorphism between TM and T^*M , where M is some Riemannian manifold, so the skew-symmetric bilinear form Ω induces a natural isomorphism between V and V^* .

Proposition 2.5. *Given a symplectic vector space (V, Ω) , the non-degenerate bilinear form Ω induces an isomorphism between V and V^* through the map*

$$V \rightarrow V^* \quad v \mapsto \Omega(v, \cdot)$$

Proof. The non-degeneracy of Ω shows this map is injective, while we know that $\dim V = \dim V^*$. Hence this is indeed an isomorphism. \square

Definition 2.6. Let (V, Ω) be a finite dimensional symplectic vector space and $Y \subset V$ a subspace. The **symplectic complement** of Y is defined to be the subspace

$$Y^\Omega := \{v \in V ; \Omega(v, u) = 0 \text{ for all } u \in Y\}.$$

For a subspace $Y \subset V$, consider the map

$$\Phi : V \rightarrow Y^* \quad v \mapsto \Omega(v, \cdot)|_Y$$

It's clear that $\ker \Phi = Y^\Omega$. The surjectivity of Φ follows by combining Proposition 2.5 together with the fact that any element of $\alpha \in Y^*$ can be extended to an element of $\tilde{\alpha} \in V^*$ such that $\tilde{\alpha}|_Y = \alpha$. It follows, by the first isomorphism theorem, $V/Y^\Omega \cong Y^*$. Since $\dim Y = \dim Y^*$, we have that $\dim V = \dim Y + \dim Y^\Omega$. Moreover, by definition, $\Omega|_{Y \times Y}$ is non-degenerate if and only if $Y \cap Y^\Omega = 0$. That is $\Omega|_{Y \times Y}$ is non-degenerate if and only if $V = Y \oplus Y^\Omega$. This leads to the following definition.

Definition 2.7. If $\Omega|_{Y \times Y} \equiv 0$ then Y is called an **isotropic** subspace of V . If Y is isotropic and $\dim Y = \frac{1}{2} \dim V$ then Y is called a **Lagrangian subspace** of V .

The above remarks give us

Proposition 2.8. *A subspace $Y \subset V$ is Lagrangian if and only if $Y = Y^\Omega$.*

2.2 Symplectic Manifolds

Let M be a manifold and let $\omega \in \Omega^2(M)$ be a 2-form. By definition, for each $p \in M$ we have that $\omega(p) := \omega_p$ is a skew-symmetric bilinear map $\omega_p : T_p M \times T_p M \rightarrow \mathbb{R}$.

Definition 2.9. A 2-form $\omega \in \Omega^2(M)$ is said to be **symplectic** if it is closed and if ω_p is symplectic (non-degenerate) for each $p \in M$. In such a case, the pair (M, ω) is called a **symplectic manifold**. By the standard form theorem, a symplectic manifold is necessarily even dimensional.

Definition 2.10. Given a 1-form $\mu \in T^*M$, the unique vector field V_μ in T^*M such that $\omega(V_\mu, \cdot) = \mu$ is called the **symplectic dual** of μ . That is, for each $p \in M$ we set $V_\mu(p)$ to be the pre-image of $\mu(p)$ under the map defined in Proposition 2.5. In other words, V_μ is the unique vector field satisfying

$$V_\mu \lrcorner \omega = \mu$$

Example 2.11. (Prototype of a Symplectic Manifold)

Let $M = \mathbb{R}^{2n}$ with standard coordinates $x^1, \dots, x^n, y_1, \dots, y_n$. The form $\omega_0 := dx^i \wedge dy_i$ is symplectic, and $T_p M \cong \mathbb{R}^{2n}$ has symplectic basis $\left\{ \frac{\partial}{\partial x^1} \Big|_p, \dots, \frac{\partial}{\partial x^n} \Big|_p, \frac{\partial}{\partial y_1} \Big|_p, \dots, \frac{\partial}{\partial y_n} \Big|_p \right\}$ so that $(\mathbb{R}^{2n}, \omega_0)$ is a symplectic manifold.

The Darboux theorem shows why the manifold above can be thought of as the prototype of symplectic manifolds. The theorem locally classifies symplectic manifolds up to symplectomorphism. That is, locally every symplectic manifold is symplectomorphic to $(\mathbb{R}^{2n}, \omega_0)$.

Theorem 2.12. (Darboux) Let (M, ω) be a symplectic manifold. For any $p \in M$ there exists a coordinate chart $(U, x^1, \dots, x^n, y_1, \dots, y_n)$ centred at p such that

$$\omega|_U = \sum_{i=1}^n dx^i \wedge dy_i$$

The coordinates giving this local expression of ω are called **Darboux coordinates**.

Proof. The proof is just an application of the Frobenius theorem together with a characterization of Darboux coordinates. See [4], page 349 for the details. \square

In the same way we defined Lagrangian subspaces of a vector space, we can define Lagrangian submanifolds.

Definition 2.13. Given a symplectic manifold (M, ω) , a submanifold (N, ι) of M is called a **Lagrangian submanifold** if at each $p \in N$, $T_p N$ is a Lagrangian subspace of $T_p M$. That is, N is Lagrangian if and only if $\iota^* \omega = \omega|_{T_p N \times T_p N} = 0$ and $\dim N = \frac{1}{2} \dim M$.

We finish this subsection with a simple yet important proposition.

Proposition 2.14. Let (M_1, ω_1) and (M_2, ω_2) be symplectic manifolds. If (L, ι) is a Lagrangian submanifold of (M_1, ω_1) and $f : (M_1, \omega_1) \rightarrow (M_2, \omega_2)$ is a symplectomorphism, then $(f(L), f \circ \iota)$ is a Lagrangian submanifold of (M_2, ω_2) .

Proof. By definition, we have that $f^*\omega_2 = \omega_1$. Hence

$$(f \circ \iota)^*\omega_2 = \iota^*f^*\omega_2 = \iota^*\omega_1 = 0$$

since (L, ι) is a Lagrangian submanifold of (X_1, ω_1) . \square

Using the results from this section we can answer the question of when a diffeomorphism between two symplectic manifolds is a symplectomorphism.

2.3 When is a Diffeomorphism a Symplectomorphism?

Let $\varphi : (M_1, \omega_1) \rightarrow (M_2, \omega_2)$ be a diffeomorphism of two symplectic manifolds. We will see that the answer to the posed question of this subsection is “if and only the graph of φ is a Lagrangian submanifold of the ‘twisted’ symplectic manifold $(M_1 \times M_2, \tilde{\omega})$.” We first formalize the definitions in this statement.

Given the two symplectic manifolds (M_1, ω_1) and (M_2, ω_2) as above, consider their Cartesian product $M_1 \times M_2$. Let π_1 and π_2 denote the projection maps onto the first and second factors respectively. For any $a, b \in \mathbb{R} \setminus \{0\}$, consider the 2-form

$$\omega := a(\pi_1^*\omega_1) + b(\pi_2^*\omega_2)$$

Since the exterior derivative commutes with the pull-back, it follows ω is closed. Moreover, to see that ω is symplectic, let $(p, q) \in M_1 \times M_2$ be arbitrary and consider non-zero $(V_p, W_q) \in T_p M_1 \times T_q M_2$. Without loss of generality, suppose that V_p is nonzero. By the non-degeneracy of ω_1 there exists $X_p \in T_p M$ such that $(\omega_1)_p(V_p, X_p) \neq 0$ so that $\omega((p, V_p), (q, 0_q)) = a \cdot \omega_{1,p}(V_p, X_p) \neq 0$.

Definition 2.15. In particular, taking $a = 1$ and $b = -1$ we obtain the **twisted product** symplectic form $\tilde{\omega} \in \Omega^2(M_1 \times M_2)$:

$$\tilde{\omega} = \pi_1^*\omega_1 - \pi_2^*\omega_2$$

Let $\Gamma_\varphi = \{(p, \varphi(p)); p \in M\}$ denote the graph of φ . It’s clear the the function

$$f : M_1 \rightarrow \Gamma_\varphi \quad p \mapsto (p, \varphi(p))$$

is an embedding. Since Γ_φ is the image of M_1 under f , it follows that Γ_φ is a submanifold of $M_1 \times M_2$ of dimension $4n - 2n = 2n$.

Hence Γ_φ always satisfies ‘half’ of the requirements of being Lagrangian. We can now prove the statement posed at the beginning of this section.

Proposition 2.16. *The diffeomorphism φ is a symplectomorphism $\iff \Gamma_\varphi$ is a Lagrangian submanifold of $(M_1 \times M_2, \tilde{\omega})$.*

Proof. We already know that (Γ_φ, ι) is a submanifold of $(M_1 \times M_2, \tilde{\omega})$, where $\iota : \Gamma_\varphi \rightarrow M_1 \times M_2$ is the inclusion map. Let f be as above. We have that

$$\begin{aligned} \Gamma_\varphi \text{ is Lagrangian} &\iff \iota^* \tilde{\omega} = 0 \\ &\iff f^* \iota^* \tilde{\omega} = 0 && \text{since } f \text{ is a diffeomorphism} \\ &\iff (\iota \circ f)^* \tilde{\omega} = 0 \end{aligned}$$

But

$$(\iota \circ f)^* \tilde{\omega} := ((\iota \circ f)^* \circ \pi_1^*) \omega_1 - ((\iota \circ f)^* \circ \pi_2^*) \omega_2 = (\pi_1 \circ \iota \circ f)^* \omega_1 - (\pi_2 \circ \iota \circ f)^* \omega_2 = \omega_1 - \varphi^* \omega_2$$

Hence

$$\Gamma_\varphi \text{ is Lagrangian} \iff \varphi^* \omega_2 = \omega_1$$

□

Remark 2.17. It is crucial in the above proof that the 2-form on $M_1 \times M_2$ is the twisted product form, otherwise this would not work.

2.4 Canonical Symplectic Structure of Cotangent Bundles

Given an arbitrary manifold M , the total space of the cotangent bundle T^*M can always be turned into a symplectic manifold. This subsection describes how.

Let M be an arbitrary n -dimensional manifold and T^*M the cotangent bundle. To turn T^*M into a symplectic manifold we need to find a closed symplectic 2-form $\omega \in \Omega^2(T^*M)$. Consider first the 1-form $\alpha \in \Omega^1(T^*M)$ defined by

$$\alpha_{(p, \xi_p)}(V_{(p, \xi_p)}) := \xi_p(\pi_*(V_{(p, \xi_p)}))$$

where $(p, \xi_p) \in T_p^*M$ and $V_{(p, \xi_p)} \in T_{(p, \xi_p)}(T^*M)$ are arbitrary and π_* is the differential of the projection map $\pi : T^*M \rightarrow M$. Define $\omega := -d\alpha$. By definition, both α and ω are global forms on T^*M . After the computation of α and ω in local coordinates, shown below, it is straightforward to verify that ω is symplectic. It is clear that ω is closed, since it is exact. Hence (T^*M, ω) is a symplectic manifold.

Definition 2.18. The 1-form $\alpha \in \Omega^1(T^*M)$ is called the **tautological 1-form** and the 2-form $\omega \in \Omega^2(T^*M)$ is called the **canonical symplectic 2-form**.

For future use we compute here α and ω in local coordinates. Let (U, x^1, \dots, x^n) be an arbitrary coordinate chart in M and $(T^*U, x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ the induced chart on T^*M . The first thing to show is how $\pi_* : T(T^*M) \rightarrow TM$ works. For arbitrary $(p, \sigma_p) \in T_p^*M$ and $W_{(p, \sigma_p)} = W^i \frac{\partial}{\partial x^i} \Big|_{(p, \sigma_p)} +$

$\widetilde{W}^i \frac{\partial}{\partial \xi_i} \Big|_{(p, \sigma_p)} \in T_{(p, \sigma_p)}(T^*M)$ there exists $a^i \in \mathbb{R}$ such that $\pi_*(W_{(p, \sigma_p)}) = a^i \frac{\partial}{\partial x^i} \Big|_p$. It follows

$$\begin{aligned} a^i &= (\pi_*(W_{(p, \sigma_p)}))(dx^i|_p) \\ &= W_{(p, \sigma_p)}(x^i \circ \pi) \\ &= W^j \frac{\partial}{\partial x^j} \Big|_{(p, \sigma_p)} (x^i \circ \pi) + \widetilde{W}^j \frac{\partial}{\partial \xi_j} \Big|_{(p, \sigma_p)} (x^i \circ \pi) \\ &= W^i \end{aligned}$$

That is,

$$\pi_*(W_{(p, \sigma_p)}) = \left(p, W^i \frac{\partial}{\partial x^i} \Big|_p \right).$$

Since α is an element of $\Gamma(T^*(T^*M))$ we have functions $a_1, \dots, a_n, b^1, \dots, b^n \in C^\infty(\pi^{-1}(U))$ such that $\alpha = a_i dx^i + b^i d\xi_i$. By definition, for arbitrary $(p, \sigma_p) \in T_p^*M$

$$a_i(p, \sigma_p) = \alpha_{(p, \sigma_p)} \left(\frac{\partial}{\partial x^i} \Big|_{(p, \sigma_p)} \right) = \sigma_p \left(\pi_* \left(\frac{\partial}{\partial x^i} \Big|_{(p, \sigma_p)} \right) \right) = \sigma_p \left(\frac{\partial}{\partial x^i} \Big|_p \right) = \sigma_i(p) = \xi_i(p, \sigma_p)$$

and

$$b^i(p, \sigma_p) = \alpha_{(p, \sigma_p)} \left(\frac{\partial}{\partial \xi_i} \Big|_{(p, \sigma_p)} \right) = \sigma_p \left(\pi_* \left(\frac{\partial}{\partial \xi_i} \Big|_{(p, \sigma_p)} \right) \right) = \sigma_p(p, 0_p) = 0 = 0(p, 0_p).$$

We have shown that in local coordinates

$$\alpha = \xi_i dx^i$$

and it follows

$$\omega := -d\alpha = dx^i \wedge d\xi_i$$

Given another 1-form $\mu \in \Omega^1(M)$ we now show that the graph of μ , considered as a function $M \rightarrow T^*M$, is a Lagrangian submanifold of T^*M if and only if μ is closed. To avoid confusion, let s_μ denote the map $s_\mu : M \rightarrow T^*M$ given by $p \mapsto (p, \mu_p)$ and let Γ_{s_μ} denote the image of μ in T^*M

$$s_\mu(M) := \Gamma_{s_\mu} = \{(p, \mu_p) ; p \in M\}$$

That is, the image of μ as a map is the same thing as the graph of s_μ . Let $\pi : T^*M \rightarrow M$ denote the projection mapping. It's clear that $\pi \circ s_\mu = \text{id}$.

Proposition 2.19. *Let α be the tautological 1-form on T^*M . Then $s_\mu^* \alpha = \mu$.*

Proof. Fix arbitrary $(p, \mu_p) \in \Gamma_{s_\mu}$. By definition, $\alpha_{(p, \mu_p)}(V) = \mu_p(\pi_* V)$. Hence for arbitrary $V \in T_p M$,

$$s_\mu^* \alpha(V) = \alpha((s_\mu)_* V) = \mu_p(\pi_* (s_\mu)_* V) = \mu_p((\pi \circ s_\mu)_* V) = \mu_p(\text{id}(V)) = \mu_p(V).$$

□

Using this we get

Proposition 2.20. Γ_{s_μ} is a Lagrangian submanifold of $T^*M \iff \mu$ is closed.

Proof. Let $\tau : M \rightarrow \Gamma_{s_\mu}$ be the same map as s_μ but with range restricted to Γ_{s_μ} . It follows that τ is a diffeomorphism and $s_\mu = \iota \circ \tau$. Hence

$$\begin{aligned} \Gamma_{s_\mu} \text{ is Lagrangian} &\iff \iota^* \omega \equiv 0 \\ &\iff \iota^* d\alpha \equiv 0 \\ &\iff \tau^* \iota^* d\alpha \equiv 0 && \text{because } \tau \text{ is a diffeomorphism} \\ &\iff (\iota \circ \tau)^* d\alpha \equiv 0 \\ &\iff (s_\mu)^* d\alpha \equiv 0 \\ &\iff d(s_\mu)^* \alpha \equiv 0 \\ &\iff d\mu \equiv 0 \\ &\iff \mu \text{ is closed} \end{aligned}$$

□

2.5 Lifting a Diffeomorphism

Definition 2.21. Given a diffeomorphism $f : M_1 \rightarrow M_2$ between two manifolds M_1 and M_2 , there is an induced symplectomorphism $f_\sharp : T^*M_1 \rightarrow T^*M_2$ called the **lift** of f which is constructed as follows.

Since f is a diffeomorphism we have that $f^* : T^*M \rightarrow T^*M$ is an isomorphism. For arbitrary $(p_1, \xi_{p_1}) \in T_{p_1}^* M_1$ we define f_\sharp by

$$f_\sharp(p_1, \xi_{p_1}) := (f(p_1), (f^*)^{-1}(\xi_{p_1})).$$

Since f is a diffeomorphism we get that both f_\sharp and f_\sharp^{-1} are bijective and smooth. Moreover, we

have the following commutative diagram.

$$\begin{array}{ccc}
T^*M_1 & \xrightarrow{f_{\sharp}} & T^*M_2 \\
\pi_1 \downarrow & & \downarrow \pi_2 \\
M_1 & \xrightarrow{f} & M_2
\end{array} \quad (\dagger)$$

Proposition 2.22. *Let α_1 and α_2 denote the tautological forms on T^*M_1 and T^*M_2 respectively. Then*

$$f_{\sharp}^*(\alpha_2) = \alpha_1.$$

Proof. Let $(p_1, \xi_{p_1}) \in T_{p_1}^*M_1$ and $(p_2, \xi_{p_2}) \in T_{p_2}^*M_2$ be such that $p_2 = f(p_1)$ and $\xi_{p_2} = f^*\xi_{p_1}$. It needs to be shown that $(f_{\sharp}^*(\alpha_2))_{(p_2, \xi_{p_2})} = (\alpha_1)_{(p_1, \xi_{p_1})}$. By definition, $(f_{\sharp}^*(\alpha_2))_{(p_2, \xi_{p_2})} \in T_{(p_1, \xi_{p_1})}^*(T^*M_1)$ so let $\eta \in T_{(p_1, \xi_{p_1})}(T^*M_1)$ be arbitrary. Then

$$\begin{aligned}
f_{\sharp}^* \left((\alpha_2)_{(p_2, \xi_{p_2})} \right) (\eta) &:= (\alpha_2)_{(p_2, \xi_{p_2})} \left(((f_{\sharp})_*) \eta \right) \\
&:= \xi_{p_2} \circ (\pi_2)_* \left(f_{\sharp*} \eta \right) \\
&= \xi_{p_2} \left((\pi_2 \circ f_{\sharp})_* \eta \right) \\
&= \xi_{p_2} \left((f \circ \pi_1)_* \eta \right) && \text{by } \dagger \\
&= \xi_{p_2} (f_* (\pi_{1*} \eta)) \\
&= f^* \xi_{p_2} (\pi_{1*} \eta) \\
&= \xi_{p_1} (\pi_{1*} \eta) \\
&= (\alpha_1)_{(p_1, \xi_{p_1})} \eta
\end{aligned}$$

□

Corollary 2.23. *In the setup of Proposition 2.22, if we take $M_1 = M_2 = M$ and let $f : M \rightarrow M$ be a diffeomorphism, then the lift of f preserves ω . That is, $f_{\sharp}^*(\omega) = \omega$.*

Proof. This follows immediately from the fact that the pull back commutes with the exterior derivative. □

The following Lemma and Theorem will be needed in section 6 to study Noether's theorem.

Lemma 2.24. Let M be a manifold. Fix $X \in \Gamma(TM)$ and let θ_t denote its flow. There exists a unique vector field X_{\sharp} on the cotangent bundle (i.e. $X_{\sharp} \in \Gamma(T(T^*M))$) such that the flow of X_{\sharp} , say Ψ_t , is the lift of θ_t . That is, $\Psi_t = \theta_{t, \sharp}$. Note that by Corollary 2.23, each Ψ_t is a symplectomorphism.

Proof. Let θ_t denote the flow of $X \in \Gamma(TM)$. We have that θ_t is a diffeomorphism $\theta_t : M \rightarrow M$ so its lift $\theta_{t, \sharp}$ is a symplectomorphism $\theta_{t, \sharp} : T^*M \rightarrow T^*M$. Proposition 2.22 shows that $\theta_{t, \sharp}$ preserves

α . Just let $X_\#$ be the infinitesimal generator of $\theta_{t,\#}$. Here the integral curves are of the form $\theta_\#^{(p,\xi)} : \mathbb{R} \rightarrow T^*M$, $t \mapsto \theta_{t,\#}(p,\xi)$, and so $\theta_{t,\#}$ is a local flow of $X_\#$. \square

Theorem 2.25. (Lifting to the Cotangent Bundle)

Let M be a manifold. Let $\alpha \in \Gamma(T^*(T^*M))$ denote the tautological 1-form on T^*M and consider the symplectic manifold $(T^*M, \omega = -d\alpha)$. If $g : T^*M \rightarrow T^*M$ is a symplectomorphism preserving α (i.e. $g^*\alpha = \alpha$) then there exists a diffeomorphism $f : M \rightarrow M$ such that $g = f_\#$.

Proof. The proof of this theorem is done by combining the following claims. For what is below we let V denote the symplectic dual of α . That is, $\omega(V, \cdot) = V \lrcorner \omega = \alpha$.

Claim 2.26. If $g^*\alpha = \alpha$ then g commutes with the flow of V , or equivalently $g_*V = V$.

Proof. Let θ_t denote the flow of V . It needs to be shown that $g \circ \theta_t = \theta_t \circ g$, or equivalently, that $g \circ \theta_t \circ g^{-1} = \theta_t$. By definition, for each $p \in M$, we have that $\theta^{(p)}$ is the unique curve satisfying $\theta^{(p)}(0) = p$ and $\frac{d}{dt}\big|_{t=0} \theta^{(p)} = V_p$. Since θ_0 is the identity, we have that $g \circ \theta_0 \circ g^{-1}(p) = p$. Hence, by uniqueness, it suffices to show that $\frac{d}{dt}\big|_{t=0} g \circ \theta_t \circ g^{-1}(p) = V_p$ which happens, by the non-degeneracy of ω , if and only if

$$\left(\frac{d}{dt} \bigg|_{t=0} g \circ \theta_t \circ g^{-1}(p) \right) \lrcorner \omega_p = V_p \lrcorner \omega_p = \alpha_p \quad (2.1)$$

By the chain rule

$$\begin{aligned} \frac{d}{dt} \bigg|_{t=0} g \circ \theta_t \circ g^{-1}(p) &= g_{*,g^{-1}(p)} \left(\frac{d}{dt} (\theta_t(g^{-1}(p))) \right) \\ &= g_{*,g^{-1}(p)} (V_{g^{-1}(p)}) \end{aligned}$$

Fix an arbitrary $Y_p \in \Gamma(T_p M)$ and plug it into both sides of (2.1). The right hand side is

$$V_p \lrcorner \omega_p(Y_p) = \alpha_p(Y_p)$$

while the left hand side becomes

$$\begin{aligned} \left(\frac{d}{dt} \bigg|_{t=0} g \circ \theta_t \circ g^{-1}(p) \lrcorner \omega_p \right) (Y_p) &= \omega_p (g_{*,g^{-1}(p)} (V_{g^{-1}(p)}), Y_p) \\ &= \omega_p (g_{*,g^{-1}(p)} (V_{g^{-1}(p)}), g_{*,g^{-1}(p)} \circ g_{*,p}^{-1} (Y_p)) \\ &= (g^* \omega)_{g^{-1}(p)} (V_{g^{-1}(p)}, g_{*,p}^{-1} (Y_p)) \\ &= \omega_{g^{-1}(p)} (V_{g^{-1}(p)}, g_{*,p}^{-1} (Y_p)) \\ &= \alpha_{g^{-1}(p)} (g_{*,p}^{-1} (Y_p)) \\ &= (g^* \alpha)_p ((g_{*,p})^{-1} (Y_p)) \\ &= \alpha_p (Y_p) \end{aligned}$$

\square

Notice that this claim had nothing to do with the fact the symplectic manifold was a cotangent bundle. This result holds for any symplectic manifold (X, ω) for which $\omega = -d\alpha$ for some 1-form α and $g^*\alpha = \alpha$.

Claim 2.27. The integral curves, $\gamma : \mathbb{R} \rightarrow T^*M$, of V are of the form

$$\gamma^{(p, \sigma)}(t) = (p, \sigma e^{-t})$$

where $(p, \sigma) \in T^*M$ is arbitrary.

Proof. In local coordinates we know that $\alpha = \xi_i dx^i$ and $\omega = dx^i \wedge d\xi_i$. Let $V = a^i \frac{\partial}{\partial x^i} + b_i \frac{\partial}{\partial \xi_i}$ where $a^i, b_i \in C^\infty(T^*U)$. By definition

$$\alpha = V \lrcorner \omega = a^i d\xi_i - b_i dx^i$$

and so $a^i = 0$ and $b_i = -\xi_i$. Let $(p, \sigma) \in T^*M$ be arbitrary and suppose $\gamma : \mathbb{R} \rightarrow T^*M$ is an integral curve of V starting at (p, σ) . We can write $\gamma(t) = (q(t), r(t))$ and it follows

$$\gamma'(t) = V_{\gamma(t)} = (q^i(t))' \frac{\partial}{\partial x^i} \Big|_{\gamma(t)} + (r_i(t))' \frac{\partial}{\partial \xi_i} \Big|_{\gamma(t)}$$

It follows that for all $t \in \mathbb{R}$, $(q_i(t))' = 0$ and $(r_i(t))' = -\xi_i(\gamma(t)) = -r_i(t)$. That is, $q^i(t)$ is a constant function while $r_i(t) = r_i(0)e^{-t}$. By assumption $\gamma(0) = (p, \sigma)$ and so it follows $\gamma(t) = (p, \sigma e^{-t})$ □

It immediately follows that θ_t is fibre preserving. That is, $\theta_t(T_x^*M) = T_x^*M$. Also, it implies that if $g(p, \xi) = (q, \eta)$ then for all $\lambda > 0$, $g(p, \lambda\xi) = g(q, \lambda\eta)$. This is because the flow of V is complete and e^{-t} is surjective onto $(0, \infty)$. Also, by the continuity of g and θ_t we have that

$$g(p, 0_p) = g(p, \lim_{t \rightarrow \infty} \theta_t(\xi)) = \lim_{t \rightarrow \infty} g(\theta_t(p, \xi)) = \lim_{t \rightarrow \infty} \theta_t(g(p, \xi)) = g(q, 0_q)$$

Hence

$$g(p, \xi) = (q, \eta) \implies g(p, \lambda\xi) = (q, \lambda\eta) \text{ for all } \lambda \geq 0$$

Supposing that $g(p, \xi) = (q, \eta)$ consider another arbitrary element (p, σ_p) of T_p^*M . Suppose that $g(p, \sigma_p) = (\tilde{q}, \mu_{\tilde{q}})$. Then by applying the above it follows $g(p, \lambda\sigma_p) = (\tilde{q}, \lambda\mu_{\tilde{q}})$ for $\lambda = 0$. That is, $g(p, 0) = (\tilde{q}, 0)$. But $g(p, 0) = (q, 0)$ and so it must be that $q = \tilde{q}$. Hence g maps fibres to fibres.

We are now ready to construct $f : M \rightarrow M$ such that $f_\sharp = g$. Indeed, define f by

$$f : M \rightarrow M \quad p \mapsto \pi \circ g(p, 0_p).$$

Claim 2.27 shows that f is well defined, while it readily follows that

$$f \circ \pi = \pi \circ g. \quad (2.2)$$

To prove $f_{\sharp} = g$, we will show $H := g \circ f_{\sharp}^{-1}$ is the identity map. By definition, H is a map from T^*M to T^*M . Let $(p, \sigma_p) \in T^*M$ be arbitrary and suppose that $H(p, \sigma_p) = (q, \eta_q)$. Let $V_{(p, \sigma_p)} \in \Gamma(T_{(p, \sigma_p)}(T^*M))$ be arbitrary. By hypothesis, g preserves α while Proposition 2.22 shows that f_{\sharp} also preserves α . Hence H preserves α . That is,

$$\begin{aligned} (H^* \alpha_{(q, \eta_q)})(V_{(p, \sigma_p)}) &= (H^* \alpha)_{(p, \sigma_p)}(V_{(p, \sigma_p)}) \\ &= \alpha_{(p, \sigma_p)}(V_{(p, \sigma_p)}) \\ &:= \sigma_p(\pi_*(V_{(p, \sigma_p)})) \end{aligned}$$

On the other hand,

$$\begin{aligned} (H^* \alpha_{(q, \eta_q)})(V_{(p, \sigma_p)}) &= \alpha_{(q, \eta_q)}(H_*(V_{(p, \sigma_p)})) \\ &= \eta_q(\pi_*(H_*(V_{(p, \sigma_p)}))) \\ &= \eta_q((\pi \circ g \circ f_{\sharp}^{-1})_*(V_{(p, \sigma_p)})) \\ &= \eta_q((f \circ \pi \circ f_{\sharp}^{-1})_*(V_{(p, \sigma_p)})) && \text{by (2.2)} \\ &= \eta_q(\pi_*(V_{(p, \sigma_p)})) && \text{by } \dagger \end{aligned}$$

But if $\sigma_p(\pi_*(V_{(p, \sigma_p)})) = \eta_q(\pi_*(V_{(p, \sigma_p)}))$ for all $V_{(p, \sigma_p)} \in T_{(p, \sigma_p)}(T^*M)$ then it must be that $(p, \sigma_p) = (q, \eta_q) = H(p, \sigma_p)$. That is, H is the identity map. \square

2.6 Constructing Symplectomorphisms

Given two manifolds M_1 and M_2 , we demonstrated in Section 2.4 that their cotangent bundles have a canonical symplectic structure. Let α_1 and α_2 denote the tautological 1-forms on T^*M_1 and T^*M_2 respectively. Let $\omega_1 = -d\alpha_1$ and $\omega_2 = -d\alpha_2$ denote the canonical 2-forms on T^*M_1 and T^*M_2 respectively. A straightforward calculation shows that the tautological 1-form on $T^*M_1 \times T^*M_2$ is

$$\alpha = \pi_1^* \alpha_1 + \pi_2^* \alpha_2.$$

implying that the canonical symplectic form on $(T^*M_1 \times T^*M_2) \cong T^*(M_1 \times M_2)$ is

$$\omega = \pi_1^* \omega_1 + \pi_2^* \omega_2.$$

By Proposition 2.16 we know that if the graph of a diffeomorphism $\varphi : (T^*M_1, \omega_1) \rightarrow (T^*M_2, \omega_2)$ is a Lagrangian submanifold of the ‘twisted product’ $(T^*M_1 \times T^*M_2, \tilde{\omega})$, then φ is a symplectomorphism. While by Proposition 2.20 we know that the graph of a 1-form $\mu \in \Omega^1(M_1 \times M_2)$ is a Lagrangian submanifold of $T^*(M_1 \times M_2)$ if the 1-form is closed. Hence, in particular, we have that

Γ_{df} is a Lagrangian submanifold of $(T^*M_1, \omega_1) \times (T^*M_2, \omega_2) \cong (T^*(M_1 \times M_2), \omega)$ for any smooth function $f \in C^\infty(M_1 \times M_2)$. Also, by Proposition 2.14 we know that symplectomorphisms take Lagrangian submanifolds to Lagrangian submanifolds. With this in mind, we first find a symplectomorphism from $(T^*(M_1 \times M_2), \omega)$ to $(T^*(M_1 \times M_2), \tilde{\omega})$ so that we can find Lagrangian submanifolds of the later. After doing this, we ‘construct’ a symplectomorphism between T^*M_1 and T^*M_2 by finding a diffeomorphism $\tau : T^*M_1 \rightarrow T^*M_2$ whose graph, which is a subset of $T^*(M_1 \times M_2)$, equals the graph of df . We will see that the existence of such a diffeomorphism is governed by the implicit function theorem.

Consider the functions

$$\sigma_2 : T^*M_2 \rightarrow T^*M_2, (p, \xi) \mapsto (p, -\xi)$$

and

$$\sigma := \text{id} \times \sigma_2 : T^*M_1 \times T^*M_2 \rightarrow T^*M_1 \times T^*M_2$$

Proposition 2.28. *The map $\sigma : (T^*M_1 \times T^*M_2, \omega) \rightarrow (T^*M_1 \times T^*M_2, \tilde{\omega})$ is a symplectomorphism. That is, $\sigma^*\tilde{\omega} = \omega$*

Proof. First note that σ is involutive and so bijective. Moreover in local coordinates $x^1, \dots, x^n, \xi_1, \dots, \xi_n$ on T^*M_2 we have

$$\sigma_2^*\alpha_2 = \sigma_2^*(\xi_i dx^i) = (\xi_i \circ \sigma_2)(d(x^i \circ \sigma_2)) = -\xi_i dx^i = -\alpha_2$$

and so

$$\sigma^*\omega = \sigma^*(\pi_1^*\omega_1) + \sigma^*(\pi_2^*\omega_2) = (\pi_1 \circ \sigma)^*\omega_1 + (\pi_2 \circ \sigma)^*\omega_2 = \pi_1^*\omega_1 - \pi_2^*\omega_2 = \tilde{\omega}$$

□

Definition 2.29. If Y is a Lagrangian submanifold of $(T^*M_1 \times T^*M_2, \omega)$ we define the **twist** of Y , denoted Y^σ to be the image of Y under σ . That is, $Y^\sigma := \sigma(Y)$.

Proposition 2.30. *If Y is a Lagrangian submanifold of $(T^*M_1 \times T^*M_2, \omega)$ then the twist of Y is a Lagrangian submanifold of $(T^*M_1 \times T^*M_2, \tilde{\omega})$*

Proof. Since $\sigma : (T^*M_1 \times T^*M_2, \omega) \rightarrow (T^*M_1 \times T^*M_2, \tilde{\omega})$ is a symplectomorphism, this result is a corollary of Proposition 2.14. □

As mentioned at the beginning of this section, we now want to find a diffeomorphism whose graph equals the graph of the closed 1-form df , where $f \in C^\infty(M_1 \times M_2)$. We call the graph of df the ‘Lagrangian submanifold generated by f ’. Before stating this formally, we introduce some notation so that we can write this submanifold in a way that will allow us to find conditions on when it is the graph of a diffeomorphism φ . By definition, we have that $(df)_{(x,y)} = \pi_1^*((d_1f)_x) + \pi_2^*((d_2f)_y)$

where x^1, \dots, x^n and y^1, \dots, y^n are local coordinates on M_1 and M_2 respectively, π_1 and π_2 are the natural projections on $T^*M_1 \times T^*M_2$ and $d_1f = \frac{\partial f}{\partial x^i} dx^i$ and $d_2f = \frac{\partial f}{\partial y^i} dy^i$.

Definition 2.31. The **Lagrangian submanifold generated by f** is the Lagrangian submanifold of $M_1 \times M_2$ defined by

$$Y_f := \{((x, y), (df)_{(x, y)}) ; (x, y) \in M_1 \times M_2\} = \{((x, y), ((d_1f)_x, (d_2f)_y)) ; (x, y) \in M_1 \times M_2\}$$

Definition 2.32. If there exists a diffeomorphism $\varphi : T^*M_1 \rightarrow T^*M_2$ such that $Y_f^\sigma = \Gamma_\varphi$ then, by Proposition 2.16, φ is a symplectomorphism. If such a symplectomorphism exists we call it the **symplectomorphism generated by f** . Recall that here Y_f^σ is the twist of Y_f .

We are trying to find a diffeomorphism $\varphi : T^*M_1 \rightarrow T^*M_2$ such that $\Gamma_\varphi = Y_f^\sigma$. But notice that

$$\begin{aligned} \Gamma_\varphi = Y_f^\sigma &\iff \{((x, \xi), (y, \eta)) : \varphi(x, \xi) = (y, \eta)\} = \{((x, (d_1f)_x), (y, -(d_2f)_y))\} \\ &\iff \xi = (d_1f)_x \text{ and } \eta = -(d_2f)_y \\ &\iff \xi_i = \frac{\partial f}{\partial x^i}(x, y) \quad (\star) \quad \text{and} \quad \eta_i = -\frac{\partial f}{\partial y^i}(x, y) \quad (\star\star) \end{aligned}$$

That is, writing $\varphi(x, \xi) = (\varphi_1(x, \xi), \varphi_2(x, \xi))$, for Γ_φ to equal Y_f^σ all we need is that $\varphi_2(x, \xi) = -(d_2f)_y$, for then it automatically follows $\varphi_1(x, \xi) = y$. Given any (x, ξ) the implicit function theorem says (locally) when a solution to (\star) exists. That is, it tells us when one can write y as a function of both x and ξ . If $(x^1, \dots, x^n, y^1, \dots, y^n)$ are local coordinates on $M_1 \times M_2$ the implicit function theorem says that we can write y as a function of x and ξ locally if and only if

$$\det \left[\frac{\partial}{\partial y^j} \left(\frac{\partial f}{\partial x^i} \right) \right] \neq 0.$$

Also, note that if we have a solution to (\star) say, $y = \varphi_1(x, \xi)$ then we can plug this solution into $(\star\star)$ and thus completely determine the map φ satisfying $\Gamma_\varphi = Y_f^\sigma$. We give some examples of this process in the next section.

2.7 Applications to Geodesic Flow

Recall the definition of geodesic flow.

Definition 2.33. Let (M, g) be a Riemannian manifold. The **geodesic flow** of M is the local \mathbb{R} -action on TM defined by

$$\Theta : \mathbb{R} \times TM \rightarrow TM, \quad (t_0, V_p) \mapsto \frac{d}{dt} \Big|_{t=t_0} \gamma_{V_p}(t)$$

where γ_{V_p} is the unique geodesic starting at $p \in M$ with initial velocity V_p .

Recall that a Riemannian manifold is called **geodesically complete** if every geodesic is defined for all $t \in \mathbb{R}$ and is called **geodesically convex** if for any two points in the manifold there exists

a minimizing geodesic connecting them. For the rest of this section, unless stated otherwise, all manifolds are assumed to be compact. Since compact metric spaces are complete, it follows from the Hopf-Rinow theorem that all of our Riemannian manifolds are geodesically complete and geodesically convex.

Example 2.34. (Free Translational Motion)

Let $M_1 = \mathbb{R}^n = M_2$ with coordinate charts $(\mathbb{R}^n, x^1, \dots, x^n)$ and $(\mathbb{R}^n, y^1, \dots, y^n)$ respectively. Endow M_1 and M_2 with the standard metric. We have the respective induced coordinate charts $(T^*\mathbb{R}^n, x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ and $(T^*\mathbb{R}^n, y^1, \dots, y^n, \eta_1, \dots, \eta_n)$. Let $f \in C^\infty(M_1 \times M_2)$ be given by

$$f(x, y) = -\frac{1}{2}d(x, y)^2 = -\frac{1}{2} \sum_{i=1}^n (x^i - y^i)^2.$$

Since the metric is assumed to be the standard one it follows that $d(x, y)$ is the usual Euclidean distance. By definition,

$$Y_f^\sigma = \left\{ \left(a, b, \frac{\partial f}{\partial x^i}(a, b) dx^i, -\frac{\partial f}{\partial y^i}(a, b) dy^i \right) \mid a, b \in \mathbb{R}^n \right\}.$$

We would like to find the symplectomorphism generated by f . That is, we would like to find a map $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that Y_f^σ equals Γ_φ . In this case (\star) is $\xi_i = \frac{\partial f}{\partial x^i}(a, b) = b^i - a^i$ and $(\star\star)$ is $\eta_i = -\frac{\partial f}{\partial y^i} = b^i - a^i$. Since $\frac{\partial f}{\partial x^i} = y^i - x^i$ we have that

$$\left[\frac{\partial}{\partial y^j} \left(\frac{\partial f}{\partial x^i} \right) \right]_{ij} = \delta_{ij}$$

so that the implicit function theorem guarantees a solution to (\star) . We have shown that

$$Y_f^\sigma = \{(a, b, b - a, b - a) \mid a, b \in \mathbb{R}^n\}.$$

For fixed $a \in \mathbb{R}^n$, the implicit function theorem has shown the existence of a function $\varphi_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n$ where for every $b \in \mathbb{R}^n$ there exists $\xi \in \mathbb{R}^n$ such that $b = \varphi_1(a, \xi)$. In this case, it is obvious that every element $\xi \in T_a^*\mathbb{R}^n = \mathbb{R}^n$ is of the form $b - a$ for some $b \in \mathbb{R}^n$. We set

$$\varphi_1(a, \xi) = \varphi_1(a, b - a) = b$$

and

$$\eta = \varphi_2(a, \xi) := -\frac{\partial f}{\partial y^i}(a, b) = \xi$$

so that $\varphi(a, \xi) = (b, \eta) = (\xi + a, \xi)$ and $\Gamma_\varphi = Y_f^\sigma$. That is, φ is the symplectomorphism generated by f . Identifying $T_p^*\mathbb{R}^n$ with $T_p\mathbb{R}^n = \{\text{vectors emanating from } p\}$ we see that φ is free translational motion.

The above example is a special case of the following.

Example 2.35. (Geodesic Flow)

Let (M, g) be an arbitrary compact Riemannian manifold. As in the previous example, let $d : M \times M \rightarrow \mathbb{R}$ denote the Riemann distance function and $f \in C^\infty(M \times M)$ be given by $f = -\frac{1}{2}d(a, b)^2$. To find the symplectomorphism generated by f we need to solve

$$(\star) \quad \xi_i = d_a f \quad \text{and} \quad (\star\star) \quad \eta_i = -d_b f.$$

Using the musical isomorphism we can identify T^*M with TM . That is, for $a \in M$, we have

$$\flat : T_a M \rightarrow T_a^* M \quad V \mapsto V^\flat := g(V, \cdot).$$

Let V and W in TM be the unique vector fields such that $V^\flat = \xi$ and $W^\flat = \eta$. Then the above equations become

$$(\star) \quad g(V, \cdot) = d_a f(\cdot) \quad \text{and} \quad (\star\star) \quad g(W, \cdot) = -d_b f(\cdot).$$

We now show that under this identification the symplectomorphism generated by f is the geodesic flow. First we need the following Lemma.

Lemma 2.36. Let $U, V \in T_a M$. Then

$$\left. \frac{d}{dt} \right|_{t=0} - \frac{1}{2} d((\exp)_a(U), (\exp)_a(tV))^2 = \langle V, U \rangle.$$

Proof. First notice that for $s \in \mathbb{R}$ small enough so that $(\exp)_a(sU)$ is contained in a geodesic ball centred at $(\exp)_a(U)$, we have that $d((\exp)_a(U), (\exp)_a(sU)) = |1 - s||U|$. Indeed, let $\gamma : \mathbb{R} \rightarrow M$ be the unique geodesic starting at a with initial velocity U . Since geodesics have constant speed it follows that the length of γ over $[s, 1]$, denoted $L(\gamma|_{[s, 1]})$, is

$$\begin{aligned} L(\gamma|_{[s, 1]}) &= \int_s^1 |\gamma'(t)| dt \\ &= |U| \int_s^1 dt \\ &= |U||1 - s| \end{aligned}$$

Since the radial geodesic from $(\exp)_a(U)$ to $(\exp)_a(sU)$ is the unique minimizing curve from $(\exp)_a(U)$ to $(\exp)_a(sU)$ (for a proof of this see [10], Proposition 6.10) it follows that

$$d((\exp)_a(U), (\exp)_a(sU)) = |1 - s||U|.$$

Next, fix $(\exp)_a(U) \in M$ and consider the function, also denoted f , defined by

$$f : M \rightarrow \mathbb{R} \quad b \mapsto f((\exp)_a(U), b).$$

Notice that given $V^\perp \in T_a M$ with $V^\perp \perp U$ the Gauss Lemma shows that there exists a curve

$\beta : \mathbb{R} \rightarrow M$ starting at a with initial velocity V^\perp such that $d(\beta(t), b)$ is constant. That is, $f(\beta(t)) = c$ implying that $f_*(\beta'(0)) = 0$. Observe that we can write $V = V^\perp + \lambda U$ for some $\lambda \in \mathbb{R}$. Letting G denote the function $G : \mathbb{R} \rightarrow M$ given by $t \mapsto (\exp)_a(tV)$ it follows that

$$\begin{aligned}
\left. \frac{d}{dt} \right|_{t=0} - \frac{1}{2} d((\exp)_a(U), (\exp)_a(tV))^2 &= \left. \frac{d}{dt} \right|_{t=0} f(G(t)) \\
&= (f_*)_{G(0)} G'(0) \\
&= (f_*)_a(V) \\
&= \lambda f_*(U) \\
&= \lambda \left. \frac{d}{ds} \right|_{s=0} - \frac{1}{2} (d((\exp)_a(sU), (\exp)_a(U)))^2 \\
&= \lambda \left. \frac{d}{ds} \right|_{s=0} - \frac{1}{2} |1-s|^2 |U|^2 \\
&= \lambda |U|^2 \\
&= \langle V, U \rangle
\end{aligned}$$

□

With this lemma we now reconsider (\star) . Evaluating the left hand side at V we get $|V|^2$. By geodesic convexity, there exists some $U \in TM$ such that $(\exp)_a(U) = b$. Using Lemma 2.36, the right hand side is

$$\begin{aligned}
d_a f(V) &= (f_*)_a(V) \\
&= \left. \frac{d}{dt} \right|_{t=0} d((\exp)_a(tV), y) \\
&= \left. \frac{d}{dt} \right|_{t=0} d((\exp)_a(tV), (\exp)_a(U)) \\
&= \langle V, U \rangle
\end{aligned}$$

Now take any vector $V' \in TM$ such that $V' \perp V$. Plugging V' into (\star) the left hand side becomes 0 and the right hand side is $\langle V', U \rangle$. Hence, $\langle V, U \rangle = \langle V, V \rangle$ and $\langle V', U \rangle = 0 = \langle V, V' \rangle$ for any $V' \perp V$. It follows that $U = V$ and so $b = (\exp)_a(V)$.

We now need to solve $(\star\star)$. We will see the solution is given by $W = \gamma'_V(1) = \left. \frac{d}{dt} \right|_{t=1} (\exp)_a(tV)$. Indeed, let $\widetilde{W} = \left. \frac{d}{dt} \right|_{t=1} (\exp)_a(tV)$ and fix any $W' \perp \widetilde{W}$. Again, by the Gauss Lemma, we have that $d_b f(W') = 0$ and so $W = k\widetilde{W}$ for some $k \in \mathbb{R}$. But since geodesics have constant speed it follows that $|V|^2 = |\widetilde{W}|^2$. Therefore the left hand side of $(\star\star)$ is

$$\langle W, \widetilde{W} \rangle = k \langle \widetilde{W}, \widetilde{W} \rangle$$

$$\begin{aligned}
&= k|\widetilde{W}|^2 \\
&= k|V|^2
\end{aligned}$$

while the right hand side is

$$\begin{aligned}
-(d_b f)(\widetilde{W}) &= -(f_*)_b(\widetilde{W}) \\
&= \left. \frac{d}{ds} \right|_{s=0} \frac{1}{2} d(a, (\exp)_a((1+s)V))^2 \\
&= \left. \frac{d}{ds} \right|_{s=0} \frac{1}{2} d((\exp)_a(0V), (\exp)_a((1+s)V))^2 \\
&= \left. \frac{d}{ds} \right|_{s=0} \frac{1}{2} (1+s^2) |V|^2 \\
&= |V|^2
\end{aligned}$$

Hence $k = 1$ showing that $W = \left. \frac{d}{dt} \right|_{t=1} (\exp)_a(V)$.

In summary, for Γ_φ to equal Y_f^σ it needs to be that V is the unique vector field such that $b = (\exp)_a(V)$ (so that b is a function of a and V) and further that $W = \left. \frac{d}{dt} \right|_{t=1} (\exp)_a(tV)$. That is, the map φ is given by

$$(a, V) \mapsto \left((\exp)_a(V), \left. \frac{d}{dt} \right|_{t=1} (\exp)_a(tV) \right) = (\gamma_V(1), \gamma'_V(1))$$

In the section on Hamiltonian mechanics, we will return to the concept of geodesic flow and give some insight as to why it arose as the symplectomorphism generated by the Riemann distance function.

3 Lagrangian Mechanics

Recall that Newton's second law states that in an inertial reference frame the motion of a particle, with position $\gamma(t)$, is given by the solution to the ODE

$$F(\gamma(t)) = p'(t)$$

where F is the net force acting on the particle and $p(t) = m(t)v(t)$ is the particle's momentum. Lagrangian mechanics is a reformulation of Newtonian mechanics in which motions are given by solutions to the Euler-Lagrange equations. In any situation where Newton's second law can be applied, so can the Euler-Lagrange equations. We will see that, in any such system, the Euler-Lagrange equations are equivalent to Newton's second law. However, the equations also hold in settings in which Newton's second law does not hold. For example, Newton's second law only holds in an inertial reference frame, while the Euler-Lagrange equations are valid in any coordinate system. Another advantage in using the Euler-Lagrange equations comes with the way in which they allow constraint forces on a mechanical system to be ignored. For example, if studying the motion of a bead on a wire, in the Lagrangian setting we do not need to worry about the forces keeping the bead constrained to the wire.

For the rest of this paper all forces are assumed to be conservative, and so we first recall this definition.

3.1 Conservative and Central Forces

Consider \mathbb{R}^n equipped with standard coordinates x^1, \dots, x^n . Let g be a metric on \mathbb{R}^n and let $x^1, \dots, x^n, v^1, \dots, v^n$ denote the induced coordinates on $T\mathbb{R}^n = \mathbb{R}^{2n}$. Recall the kinetic energy is defined by

$$K : \mathbb{R}^{2n} \rightarrow \mathbb{R} \quad (x, v) \mapsto \frac{1}{2}mg(v, v) = \frac{1}{2}mg_{ij}v^i v^j = \frac{1}{2}m|v|^2$$

where m is some positive constant, called the mass of the particle. Fix two points $r_1, r_2 \in \mathbb{R}^n$ and consider a curve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ such that $\gamma(a) = r_1$ and $\gamma(b) = r_2$. Then the work done by a force F on a particle moving along γ is defined to be the integral of F over the curve γ

$$W(r_1 \rightarrow r_2, \gamma) := \int_{\gamma} F \cdot ds = \int_a^b F(\gamma(t)) \cdot \gamma'(t) dt$$

Theorem 3.1. (Work-Kinetic Energy Theorem)

Given a system of k particles of masses m_i with position r_i , the change in kinetic energy of the system is equal to the sum of the work done on each particle.

Proof. Let F_i denote the force acting on the i -th particle and W_i the work done by F_i on the i -th

particle. By Newton's second law we have that

$$\left. \frac{d}{dt} \right|_{r(t)} K = \sum_{i=1}^n (m_i r_i''(t)) \cdot r_i'(t) = \sum_{i=1}^n F_i(r_i(t)) \cdot r_i'(t).$$

It follows that

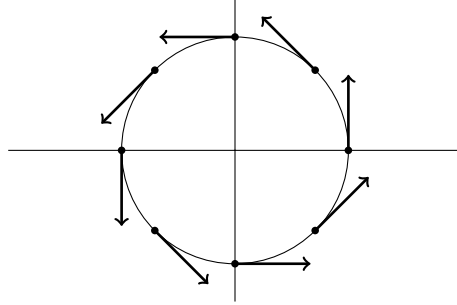
$$K(t_2) - K(t_1) = \int_{t_1}^{t_2} \frac{dK}{dt} dt = \sum_{i=1}^n \int_{t_1}^{t_2} F_i \cdot r_i' dt = \sum_{i=1}^n W_i$$

□

As the next example shows, given two paths $\alpha, \beta : [a, b] \rightarrow \mathbb{R}^n$ with $\alpha(a) = r_1 = \beta(a)$ and $\alpha(b) = r_2 = \beta(b)$, it may be that $W(r_1 \rightarrow r_2, \alpha) \neq W(r_1 \rightarrow r_2, \beta)$.

Example 3.2. (A Non-Conservative Force)

Consider a particle moving around the unit circle under the following force field.



This vector field (force) has integral curve $\gamma(t) = (\cos t, \sin t)$. It follows that $F(x, y) = (-y, x)$. Fix the points $(1, 0)$ and $(-1, 0)$. Consider the curve $\alpha : [0, \pi] \rightarrow \mathbb{R}^2$ given by $t \mapsto (\cos t, \sin t)$ and $\beta : [0, \pi] \rightarrow \mathbb{R}^2$ given by $t \mapsto (\cos t, -\sin t)$. Then α and β are both curves starting at $(1, 0)$ and ending at $(-1, 0)$; however,

$$W((1, 0) \rightarrow (-1, 0), \alpha) = \int_{\alpha} F \cdot ds = \int_0^{\pi} (-\sin t, \cos t) \cdot (-\sin t, \cos t) dt = \pi$$

while

$$W((1, 0) \rightarrow (-1, 0), \beta) = \int_{\beta} F \cdot ds = \int_0^{\pi} (\sin t, \cos t) \cdot (-\sin t, -\cos t) dt = -\pi$$

Contrary to this example, for certain forces (such as the gravitational and electrostatic forces) it is the case that W does not depend on the path traversed by the particle. This leads to the following definition.

Definition 3.3. A force is called **conservative** if the work done is path independent. Letting $r_1, r_2 \in \mathbb{R}^n$ be arbitrary, this means that for any two curves $\alpha, \beta : [a, b] \rightarrow \mathbb{R}^n$ with $\alpha(a) = r_1 = \beta(a)$

and $\alpha(b) = r_2 = \beta(b)$,

$$W(r_1 \rightarrow r_2, \alpha) = \int_{\alpha} F \cdot ds = \int_{\beta} F \cdot ds = W(r_1 \rightarrow r_2, \beta)$$

In this case the work done by F is justifiably denoted $W(r_1 \rightarrow r_2)$. A mechanical system is called **conservative** if the net force is conservative.

Recall that the gradient of a function $U : \mathbb{R}^n \rightarrow \mathbb{R}$ is defined to be

$$\nabla U := (dU)^{\sharp} = \left(g^{ji} \frac{\partial U}{\partial x^i} \right) \frac{\partial}{\partial x^j}$$

where $\sharp : T^*\mathbb{R}^n \rightarrow T\mathbb{R}^n$ is the musical isomorphism and $g^{ji} = [g^{-1}]_{ji}$. With the standard metric, the above definition reduces to the standard notion of the gradient.

Theorem 3.4. *A force F is conservative if and only if there exists a continuously differentiable function $U : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $F = -\nabla U$.*

Proof. First suppose that the work done by F is conservative. Fix a point $r_0 \in \mathbb{R}^n$ and define

$$U : \mathbb{R}^n \rightarrow \mathbb{R} \quad r \mapsto - \int_{\gamma} F \cdot ds$$

where $\gamma : [a, b] \rightarrow \mathbb{R}^n$ is an arbitrary curve with $\gamma(a) = r_0$ and $\gamma(b) = r$. By hypothesis this function is well defined. By the fundamental theorem of line integrals it follows

$$\int_{\gamma} \nabla U \cdot ds = U(r) - U(r_0) = - \int_{\gamma} F \cdot ds - U(r_0).$$

In particular, taking γ to be the curve $\gamma : [0, t] \rightarrow \mathbb{R}^n$ given by $t \mapsto (1 - t)r_0 + tr$, it follows that

$$\int_0^t \nabla(\gamma(t))(r - r_0) dt = - \int_0^t F(\gamma(t))(r - r_0) dt - U(r_0).$$

But since $U(r_0)$ is a constant, it follows from the fundamental theorem of calculus that $F = -\nabla U$.

Conversely, suppose that there exists $U : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $F = -\nabla U$. Then for arbitrary γ as above,

$$- \int_{\gamma} F \cdot ds = \int_{\gamma} \nabla U \cdot ds = U(r) - U(r_0)$$

That is, $W(r_0 \rightarrow r, \gamma) = U(r) - U(r_0)$ and so only depends on the end points r_0 and r . Hence, the work done is path independent showing F is conservative. \square

Definition 3.5. For a conservative force F , the scalar function $U : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $F = -\nabla U$ is called the **potential energy**.

The reason that a force with this property is called conservative comes from Theorem 3.7 below.

Definition 3.6. The **total energy** of a conservative system is defined to be

$$E = T + U,$$

the kinetic energy plus the potential energy.

Theorem 3.7. (Conservation of Total Energy) In a conservative mechanical system, the total energy is conserved. That is, $\frac{d}{dt}E = 0$.

Proof. Suppose that $F = -\nabla U$. For arbitrary $t_1, t_2 \in \mathbb{R}$, by the Work-Kinetic Energy Theorem

$$K(\gamma(t_2)) - K(\gamma(t_1)) = W(\gamma(t_1) \rightarrow \gamma(t_2)) = \int_{\gamma} F \cdot ds = \int_{\gamma} \nabla U \cdot ds = U(\gamma(t_1)) - U(\gamma(t_2))$$

Hence $(K + U)(\gamma(t_2)) = (K + U)(\gamma(t_1))$ and so $E := K + U$ is independent of t . \square

In section 6 we will do some interesting computations with the Laplace-Runge-Lenze vector and so we recall here the two-body central force problem.

Definition 3.8. A **central force** on a particle with position vector \vec{r} is a conservative force for which the corresponding potential energy is only a function of $\|\vec{r}\|$.

The classical example of a central force is one given by a potential of the form

$$U(\vec{r}) = -\frac{k}{\|\vec{r}\|}$$

where k is some constant. Given a system of two particles, (say \vec{r}_1 and \vec{r}_2) in a closed system, by fixing one of the particles and considering the relative position vector $\vec{r} = \vec{r}_1 - \vec{r}_2$ we can give explicit formulas for the potential energy corresponding to the gravitational and Coulomb force. The gravitational potential is given by

$$U(\vec{r}) = -\frac{Gm_1m_2}{\|\vec{r}\|}$$

and the Coulomb by

$$U(\vec{r}) = \frac{1}{4\pi\epsilon_0} \frac{q_1q_2}{\|\vec{r}\|}$$

where q_1 and q_2 are the charges of the two particles. Here G and ϵ_0 are two constants whose explicit values depend on the units being used.

Proposition 3.9. *Consider a closed system of two particles moving in \mathbb{R}^3 with the standard metric. If the particles are subject to a central force field then the force is always parallel to the relative position of the two particles.*

Proof. By definition, U is a function of only $\|\vec{r}\|$. That is $U = U(\|\vec{r}\|)$. In spherical coordinates we have that $\nabla U = \frac{\partial U}{\partial r} \frac{\partial}{\partial r}$ where $\frac{\partial}{\partial r}$ is associated with the vector $\hat{r} = \frac{\vec{r}}{\|\vec{r}\|}$. \square

Definition 3.10. If \vec{r} is the position vector of a particle, then its **angular momentum** is defined to be

$$\vec{L} := \vec{r} \times p = m(\vec{r} \times \dot{\vec{r}})$$

and the **torque** is

$$\tau := \vec{r} \times F$$

Using Newton's second law, it immediately follows that torque is the time derivative of angular momentum.

Proposition 3.11. *If two particles are subject to a central force field then the angular momentum of their relative position vector is constant.*

Proof. By proposition 3.9 the force is parallel to \vec{r} . Hence $\frac{d}{dt}\vec{L} = m(\dot{\vec{r}} \times \dot{\vec{r}}) + (\vec{r} \times F) = 0$. \square

Remark 3.12. By definition, the angular momentum is always orthogonal to the momentum and position vector. In the two-body central force problem, Proposition 3.11 showed that the angular momentum vector is constant. Hence it must be that the plane determined by the momentum and position of the relative position vector is constant. That is, the movement of the two particles is always restricted to a plane. Under translation and rotation, it is no loss of generality to assume that the particles motion is restricted to the xy -plane.

3.2 The Calculus of Variations and the Euler-Lagrange Equations

The calculus of variations studies functionals on a given space X . In this section X will be the set of smooth curves $\gamma : [a, b] \rightarrow M$, where M is some manifold. The Euler-Lagrange equations will arise as the extreme points of a specific function, called the action. After deriving the Euler-Lagrange equations we will see how they are a generalization of Newton's second law.

Definition 3.13. Let M be an n -dimensional manifold with tangent bundle TM . Given $L \in C^\infty(TM)$ the pair (M, L) is called a **Lagrangian system** and L is called the **Lagrangian**.

Definition 3.14. A Riemannian manifold (M, g) with Lagrangian $L = K - U$ is called a **natural Lagrangian system**.

At first, it may seem that the Lagrangian defined in a natural system is random; however, we will see that in these systems the Euler-Lagrange equations are equivalent to Newton's second law. The Euler-Lagrange equations have proven to be more effective than Newton's second law in many different natural Lagrangian systems.

Definition 3.15. Given a smooth curve $\gamma : [a, b] \rightarrow M$ we get an induced curve $\tilde{\gamma}$, called the **lift** of γ , into the tangent bundle defined by

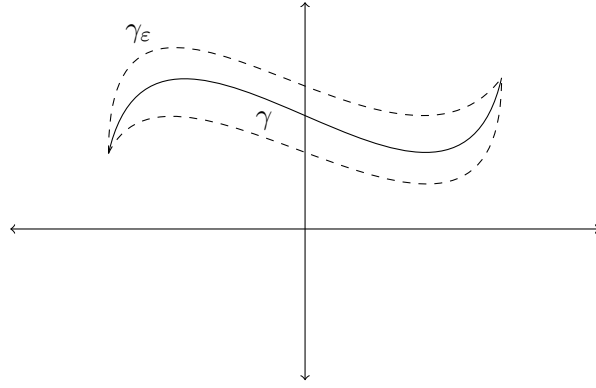
$$\tilde{\gamma} : [a, b] \rightarrow TM \quad t \mapsto (\gamma(t), \gamma'(t))$$

Definition 3.16. Let $\mathcal{C} = \{\gamma : [a, b] \rightarrow M ; \gamma \text{ is smooth}\}$. Consider the function $\mathcal{A} : \mathcal{C} \rightarrow \mathbb{R}$ defined by

$$\mathcal{A}(\gamma) := \mathcal{A}_\gamma := \int_a^b \tilde{\gamma}^* L(t) dt = \int_a^b L(\gamma(t), \gamma'(t)) dt$$

The function $\mathcal{A} : \mathcal{C} \rightarrow \mathbb{R}$ is called the **action** of the Lagrangian system.

The goal of this section is to find curves $\gamma : [a, b] \rightarrow M$ which are critical points for \mathcal{A} . Just as minimum and maximum points are extreme points in elementary calculus, we seek to find curves for which the “derivative” of the action vanishes. Intuitively, for a curve $\gamma : [a, b] \rightarrow M$ to be a minimum we need the value of \mathcal{A} to be no greater on γ than on curves ‘close’ to γ , say ‘within ε ’, as pictured below.



More precisely, fix a coordinate chart (U, x^1, \dots, x^n) in M and consider $(TU, x^1, \dots, x^n, v^1, \dots, v^n)$, the induced chart on TM . Let $\gamma : [a, b] \rightarrow U$ be a curve. Given $\varepsilon \in \mathbb{R}$, pick arbitrary $c^1, \dots, c^n \in C^\infty([a, b])$ with $c^i(a) = 0 = c^i(b)$. Define $\gamma_\varepsilon(t) := (\gamma^1(t) + \varepsilon c^1(t), \dots, \gamma^n(t) + \varepsilon c^n(t))$. Note that we can choose ε small enough so that γ_ε is contained in U . We have that $\gamma_0 = \gamma$ and so if γ is a minimum of \mathcal{A} then

$$\left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{A}_{\gamma_\varepsilon} = 0$$

On the other hand, the chain rule, product rule and fundamental theorem of calculus give that

$$\begin{aligned} \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{A}_{\gamma_\varepsilon} &= \frac{d}{d\varepsilon} \int_a^b L(\gamma_\varepsilon(t), \gamma'_\varepsilon(t)) dt \\ &= \int_a^b \left(\left(\frac{\partial L}{\partial x^i} \right)_{\tilde{\gamma}(t)} c^i(t) + \left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} \frac{d}{dt} c^i(t) \right) dt \\ &= \int_a^b \left(\left(\frac{\partial L}{\partial x^i} \right)_{\tilde{\gamma}(t)} c^i(t) + \frac{d}{dt} \left(\left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} c^i(t) \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} c^i(t) \right) dt \\ &= \int_a^b \left(\left(\frac{\partial L}{\partial x^i} \right)_{\tilde{\gamma}(t)} c^i(t) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} c^i(t) \right) dt + \left[\left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} c^i(t) \right]_a^b \\ &= \int_a^b \left(\left(\frac{\partial L}{\partial x^i} \right)_{\tilde{\gamma}(t)} c^i(t) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \right)_{\tilde{\gamma}(t)} c^i(t) \right) dt \end{aligned}$$

$$= \int_a^b \left[\left(\frac{\partial L}{\partial x^i} \Big|_{\tilde{\gamma}(t)} \right) - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \Big|_{\tilde{\gamma}(t)} \right) \right] c^i(t)$$

But since this expression is equal to zero for all $c^1(t), \dots, c^n(t)$ with $c^i(a) = c^i(b) = 0$, the *Fundamental Lemma of Calculus of Variations* (see [2] page 57) implies

$$\frac{\partial L}{\partial x^i} \Big|_{\tilde{\gamma}(t)} - \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \Big|_{\tilde{\gamma}(t)} \right) = 0 \quad \text{for all } i = 1, \dots, n$$

These n second order ODE's are called the **Euler-Lagrange** equations. Hence we have shown that a necessary condition for a curve to minimize the action is that it needs to satisfy the Euler-Lagrange equations.

Remark 3.17. If the Lagrangian is strictly convex, meaning for fixed $x \in M$, arbitrary $v, w \in T_x M$ and $0 < t < 1$ we have $L(x, tw + (1-t)v) < tL(x, w) + (1-t)L(x, v)$, then the converse is locally true. That is, if a curve $\gamma : [a, b] \rightarrow M$ satisfies the Euler-Lagrange equations, then there exists a subinterval $[a_1, b_1] \subset [a, b]$ such that $\gamma|_{[a_1, b_1]}$ minimizes the action. See [1] page 117 for a proof of this.

Remark 3.18. In the above derivation of the Euler-Lagrange equations, it was assumed that the Lagrangian was time independent. However, there are many situations in which the Lagrangian does depend on time, some of which we will see in subsequent sections. But notice that even if the Lagrangian were of the form $L(x^1, \dots, x^n, v^1, \dots, v^n, t)$ the above calculation would be exactly the same and the Euler-Lagrange equations derived above would not change.

3.3 Examples of Lagrangian Systems

Example 3.19. The Action as the Length Functional

In the case that we are working in a natural system for which the net force is zero, our Lagrangian reduces to

$$L = K : TM \rightarrow \mathbb{R} \quad (p, V_p) \mapsto \frac{1}{2} m g_p(V_p, V_p).$$

In this case $\mathcal{A}(\gamma)$ is just a constant times the length of γ . That is,

$$\mathcal{A}(\gamma) = \frac{1}{2} m \int_{\gamma} g_{\gamma(t)}(\gamma'(t), \gamma'(t)) dt.$$

A standard result from Riemannian geometry is that the critical points of the length functional are geodesics. Hence any geodesic is a critical point of \mathcal{A} . In particular, if the Riemannian manifold is \mathbb{R}^n with the standard metric, and the net force is zero, then we get that the solutions of the Euler-Lagrange equations are straight lines. That is, the shortest path between two points is a straight line. Using this mechanical system we will give, in the section on Hamiltonian mechanics, another interpretation of geodesic flow.

Example 3.20. (Natural System with Standard Coordinates)

Consider a natural Lagrangian system in \mathbb{R}^2 with the standard metric. Let x^1, x^2 denote the standard coordinates and let x^1, x^2, v^1, v^2 be the induced coordinates on $T\mathbb{R}^2 = \mathbb{R}^4$. In this coordinate system we have that $g = (dx^1)^2 + (dx^2)^2$. Suppose that a particle of mass m is moving in \mathbb{R}^2 under a conservative force field $F = -\nabla U$. By definition the Lagrangian $L : \mathbb{R}^4 \rightarrow \mathbb{R}$ is defined by

$$\begin{aligned} L(x^1, x^2, v^1, v^2) &:= K(x^1, x^2, v^1, v^2) - U(x^1, x^2, v^1, v^2) \\ &= \frac{1}{2}m(v^1)^2 + \frac{1}{2}m(v^2)^2 - U(x^1, x^2). \end{aligned}$$

The elements of the Euler-Lagrange equations are

$$\frac{\partial L}{\partial x^1}(\tilde{\gamma}(t)) = -\frac{\partial U}{\partial x^1}(\gamma(t)) \quad , \quad \frac{\partial L}{\partial x^2}(\tilde{\gamma}(t)) = -\frac{\partial U}{\partial x^2}(\gamma(t))$$

and

$$\frac{d}{dt} \left(\frac{\partial L}{\partial v^1}(\tilde{\gamma}(t)) \right) = \frac{d}{dt} (mv^1(\tilde{\gamma}(t))) = m\dot{\gamma}^1(t) \quad , \quad \frac{d}{dt} \left(\frac{\partial L}{\partial v^2}(\tilde{\gamma}(t)) \right) = \frac{d}{dt} (mv^2(\tilde{\gamma}(t))) = m\dot{\gamma}^2(t).$$

Hence, in this setting, the Euler-Lagrange equations are equivalent to Newton's second law. This equivalence easily extends to conservative force fields on \mathbb{R}^n . The same result also holds for k particles moving in \mathbb{R}^n . To see this, just take the manifold to be \mathbb{R}^{kn} so that the motion of the k -particles can be described by the motion of one particle.

Example 3.21. (Natural System with Polar Coordinates)

Consider the setup of the previous example, but with polar coordinates (r, θ) . Let $r, \theta, \tilde{r}, \tilde{\theta}$ denote the induced coordinates on $T\mathbb{R}^2 = \mathbb{R}^4$. By the chain rule we have $\frac{\partial}{\partial r} = \cos \theta \frac{\partial}{\partial x^1} + \sin \theta \frac{\partial}{\partial x^2}$ and $\frac{\partial}{\partial \theta} = -r \sin \theta \frac{\partial}{\partial x^1} + r \cos \theta \frac{\partial}{\partial x^2}$. It follows that in polar coordinates

$$g = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix} \quad \text{and} \quad g^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{r^2} \end{bmatrix}$$

In these coordinates, the velocity of $\gamma(t) = (r(t), \theta(t))$ is given by $\dot{r} \frac{\partial}{\partial r} + \dot{\theta} \frac{\partial}{\partial \theta}$. Therefore the kinetic energy of the particle at time t is

$$\begin{aligned} K(\tilde{\gamma}(t)) &= K(r, \theta, \dot{r}, \dot{\theta}) \\ &= \frac{1}{2}mg \left(\dot{r} \frac{\partial}{\partial r}, \dot{\theta} \frac{\partial}{\partial \theta} \right) \\ &= \frac{1}{2}m(\dot{r})^2 + \frac{1}{2}mr^2(\dot{\theta})^2 \end{aligned}$$

Since $L = K - U$ it follows that

$$\frac{\partial L}{\partial r} = m\dot{\theta}^2 - \frac{\partial U}{\partial r} \quad \text{and} \quad \frac{\partial L}{\partial \theta} = -\frac{\partial U}{\partial \theta}$$

while

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \Big|_{\tilde{\gamma}(t)} \right) = m\ddot{r} \quad \text{and} \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \Big|_{\tilde{\gamma}(t)} \right) = \frac{d}{dt} (mr^2\dot{\theta}) = mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta}$$

We have that $F = -\nabla U = -\frac{\partial U}{\partial r} \frac{\partial}{\partial r} - \frac{1}{r^2} \frac{\partial U}{\partial \theta} \frac{\partial}{\partial \theta}$ and by definition, the forces in the r and θ directions are

$$F_r = \frac{F \cdot \frac{\partial}{\partial r}}{\left| \frac{\partial}{\partial r} \right|} = -\frac{\partial U}{\partial r} \quad \text{and} \quad F_\theta = \frac{F \cdot \frac{\partial}{\partial \theta}}{\left| \frac{\partial}{\partial \theta} \right|} = -\frac{1}{r} \frac{\partial U}{\partial \theta}$$

Combining these equalities the first Euler Lagrange equation gives that

$$F_r = m(\ddot{r} - r\dot{\theta}^2)$$

which is the r -component of Newton's second law, while the second equation says

$$rF_\theta = -\frac{\partial U}{\partial \theta} = \frac{d}{dt} (mr^2\dot{\theta}) = mr^2\ddot{\theta} + 2mr\dot{r}\dot{\theta}$$

which is precisely the statement that the torque, rF_θ , is the derivative of angular momentum, $mr^2\dot{\theta}$.

Remark 3.22. In the above examples we showed that a motion is determined by solving the Euler-Lagrange equations. However, we do not know if this solution is a maximum or a minimum. To prove that a solution is a maximum or a minimum usually requires some extra work. However, the above examples do demonstrate **Hamilton's Principle** which is that the path a particle follows is a critical point of the action \mathcal{A} .

These examples lead to the following definition.

Definition 3.23. In a Lagrangian system (M, L) , a curve $\gamma : [a, b] \rightarrow M$ is called a **motion** if $\tilde{\gamma}(t)$ satisfies the Euler-Lagrange equations.

Since the Euler-Lagrange equations were derived from a statement about curves, independent of the coordinate system chosen, this means that if the equation holds in one coordinate system, they hold in any other. However, we can also prove this rigorously.

Proposition 3.24. *Let (M, L) be a Lagrangian system. If the Euler-Lagrange equations hold in one coordinate chart then they hold in all coordinate charts.*

Proof. Let (U, x^1, \dots, x^n) and $(V, \tilde{x}^1, \dots, \tilde{x}^n)$ denote two arbitrary coordinate charts on M with non-trivial intersection and let $x^1, \dots, x^n, v^1, \dots, v^n$ and $\tilde{x}^1, \dots, \tilde{x}^n, \tilde{v}^1, \dots, \tilde{v}^n$ denote the induced

coordinates on TU and TV respectively. Suppose that γ is a motion in (M, L) . That is

$$\left. \frac{\partial L}{\partial x^i} \right|_{\tilde{\gamma}(t)} = \frac{d}{dt} \left(\left. \frac{\partial L}{\partial v^i} \right|_{\tilde{\gamma}(t)} \right).$$

The chain rule shows that $\frac{\partial}{\partial x^i} = \frac{\partial \tilde{x}^j}{\partial x^i} \frac{\partial}{\partial \tilde{x}^j}$ implying that $\tilde{v}^i = \frac{\partial \tilde{x}^i}{\partial x^j} v^j$. Using these expressions we get

$$\frac{\partial \tilde{v}^i}{\partial v^l} = \frac{\partial \tilde{x}^i}{\partial x^j} \delta_l^j = \frac{\partial \tilde{x}^i}{\partial x^l} \quad (3.1)$$

and

$$\frac{\partial \tilde{v}^i}{\partial x^l} = \frac{\partial^2 \tilde{x}^i}{\partial x^j \partial x^l} v^j. \quad (3.2)$$

It follows that

$$\begin{aligned} \left. \frac{\partial L}{\partial x^i} \right|_{\tilde{\gamma}(t)} &= \left(\left. \frac{\partial L}{\partial \tilde{x}^j} \frac{\partial \tilde{x}^j}{\partial x^i} + \frac{\partial L}{\partial \tilde{v}^j} \frac{\partial \tilde{v}^j}{\partial x^i} \right|_{\tilde{\gamma}(t)} \right) \\ &= \left(\left. \frac{\partial L}{\partial \tilde{x}^j} \frac{\partial \tilde{v}^j}{\partial v^i} + \frac{\partial L}{\partial \tilde{v}^j} \frac{\partial^2 \tilde{x}^j}{\partial x^l \partial x^i} v^l \right|_{\tilde{\gamma}(t)} \right) && \text{by (3.1) and (3.2)} \\ &= \left(\left. \frac{\partial L}{\partial \tilde{x}^j} \frac{\partial \tilde{v}^j}{\partial v^i} + \frac{\partial L}{\partial \tilde{v}^j} \frac{\partial^2 \tilde{x}^j}{\partial x^l \partial x^i} \frac{dx^l}{dt} \right|_{\tilde{\gamma}(t)} \right) && \text{since we are substituting } \tilde{\gamma}(t). \end{aligned}$$

We also have that

$$\begin{aligned} \frac{d}{dt} \left(\left. \frac{\partial L}{\partial v^i} \right|_{\tilde{\gamma}(t)} \right) &= \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \tilde{v}^j} \frac{\partial \tilde{v}^j}{\partial v^i} \right|_{\tilde{\gamma}(t)} \right) \\ &= \frac{d}{dt} \left(\left(\left. \frac{\partial L}{\partial \tilde{v}^j} \frac{\partial \tilde{x}^j}{\partial x^i} \right|_{\tilde{\gamma}(t)} \right) \right) && \text{by (3.1)} \\ &= \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \tilde{v}^j} \right|_{\tilde{\gamma}(t)} \right) \frac{\partial \tilde{x}^j}{\partial x^i} + \left. \frac{\partial L}{\partial \tilde{v}^j} \right|_{\tilde{\gamma}(t)} \left(\frac{\partial^2 \tilde{x}^j}{\partial x^l \partial x^i} \frac{dx^l}{dt} \right) && \text{by the product rule} \\ &= \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \tilde{v}^j} \right|_{\tilde{\gamma}(t)} \right) \frac{\partial \tilde{v}^j}{\partial v^i} + \left. \frac{\partial L}{\partial \tilde{v}^j} \right|_{\tilde{\gamma}(t)} \left(\frac{\partial^2 \tilde{x}^j}{\partial x^l \partial x^i} \frac{dx^l}{dt} \right) && \text{by (3.1).} \end{aligned}$$

By combining these two calculations it follows that

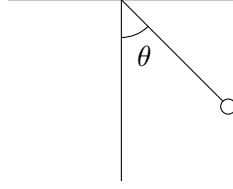
$$\left. \frac{\partial L}{\partial x^i} \right|_{\tilde{\gamma}(t)} - \frac{d}{dt} \left(\left. \frac{\partial L}{\partial v^i} \right|_{\tilde{\gamma}(t)} \right) = \left(\left. \frac{\partial L}{\partial \tilde{x}^j} \right|_{\tilde{\gamma}(t)} - \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \tilde{v}^j} \right|_{\tilde{\gamma}(t)} \right) \right) \frac{\partial \tilde{x}^j}{\partial v^i}$$

□

To make a comparison between Lagrangian and Hamiltonian mechanics, we compute here the Euler-Lagrange equations for the simple pendulum and revisit the calculation in the next section.

Example 3.25. The Simple Pendulum

The simple pendulum is illustrated below.



We have a mass m attached to a weightless rod of length l . That is, we are working in the 1-dimensional submanifold (S^1, θ) of $(\mathbb{R}^2, r, \theta)$ which we endow with the standard metric. By Newton's second law we know that the equation of motion for the mass is

$$\ddot{\theta} = -\frac{g}{l} \sin \theta$$

Let $\theta, \tilde{\theta}$ denote the induced coordinates on $T^*S^1 \cong S^1 \times \mathbb{R}$. As in example 3.21 we have that $K = \frac{1}{2}ml^2(\tilde{\theta})^2$. A standard calculation gives that the net force $F = -\frac{mg}{l} \sin \theta$ is conservative with potential energy $U = mgl(1 - \cos \theta)$. With the natural Lagrangian $L = K - U$, let $\alpha(t)$ be a motion in the Lagrangian system (S^1, L) . That is, suppose $\tilde{\alpha}(t) = (\alpha(t), \alpha'(t))$ satisfies the Euler-Lagrange equations, i.e.

$$\left. \frac{\partial L}{\partial \theta} \right|_{\tilde{\alpha}(t)} = \frac{d}{dt} \left(\left. \frac{\partial L}{\partial \tilde{\theta}} \right|_{\tilde{\alpha}(t)} \right) \implies -mgl \sin(\alpha(t)) = \frac{d}{dt} (ml^2 \alpha'(t)) = ml^2 \alpha''(t)$$

This is precisely the statement that the torque exerted by gravity on the pendulum, $-mgl \sin \alpha(t)$, is the product of the moment of inertia, ml^2 , and angular acceleration, $\alpha''(t)$.

Remark 3.26. In the above example we found the equations of motion without computing the tension of the rope, the constraint force. Although this example is simple enough to solve using Newton's second law, it gives a glimpse into how the Euler-Lagrange equations can be used to simplify other complicated systems with constraint forces.

4 Hamiltonian Mechanics

The next level of formality in studying mechanics is done using Hamilton's formulation. This approach is done through the framework of symplectic geometry, which was introduced in section 2. In section 2.4 we saw how the cotangent bundle has a canonical symplectic structure. Where Lagrangian mechanics studies curves living in TM , Hamilton's approach instead studies curves in the symplectic manifold T^*M . We will see that the outcomes predicted in Hamiltonian mechanics agree with those from Lagrangian mechanics in situations where both can be applied.

4.1 Hamiltonian Vector Fields

Definition 4.1. A triple (X, ω, H) where (X, ω) is a symplectic manifold and $H \in C^\infty(X)$ is called a **Hamiltonian system** and H is called the associated **Hamiltonian function**.

Fix a Hamiltonian system (X, ω, H) . We have that $dH \in \Gamma(T^*X)$ and by Proposition 2.5 there exists a corresponding vector field $V_H \in \Gamma(TX)$, the symplectic dual of dH . Notice that any $G \in C^\infty(X)$ induces a vector field V_G in this way.

Definition 4.2. Given $G \in C^\infty(X)$, the vector field V_G is called the **Hamiltonian vector field** associated to G .

We will need the following propositions in section 6.

Proposition 4.3. *The flow θ_t of V_H preserves ω (i.e. $\theta_t^*\omega = \omega$).*

Proof. By Cartan's magic formula together with the closedness of ω it follows that

$$\mathcal{L}_{V_H}\omega = d(V_H \lrcorner \omega) + V_H \lrcorner d\omega = d(V_H \lrcorner \omega) = d(dH) = 0$$

□

Proposition 4.4. *The Hamiltonian function of a Hamiltonian vector field is constant on its flows. That is, $H \circ \theta_t = H$ on the domain of θ_t .*

Proof. By the antisymmetry of the interior product

$$\mathcal{L}_{V_H}H = V_H H = V_H \lrcorner dH = V_H \lrcorner (V_H \lrcorner \omega) = 0$$

□

The following is an illustration of this result.

Example 4.5. (Height Function on S^2)

The sphere S^2 is a symplectic manifold when equipped with the local 2-form $d\theta \wedge dh$, where θ is a local coordinate for S^1 and h is the x^3 -coordinate of \mathbb{R}^3 . Let $H(\theta, h) = h$. It's clear that $dH = dh$ and since $\omega(V_H, \cdot) = (d\theta \wedge dh)(V_H, \cdot) = dh$ it must be that $V_H = \frac{\partial}{\partial \theta}$. The vector field $V_H = \frac{\partial}{\partial \theta}$ has flow $\rho_t : S^2 \rightarrow S^2$ given by $(\theta, h) \mapsto (\theta + t, h)$, which is clearly constant under the height function. That is, $H \circ \rho_t = H$.

By the closedness of ω we have $\mathcal{L}_W\omega = d(W \lrcorner \omega)$ for any $W \in \Gamma(TX)$. We see that the flow of an arbitrary $W \in \Gamma(TX)$ preserves ω if $W \lrcorner \omega$ is closed. This motivates the following definition.

Definition 4.6. A vector field $W \in \Gamma(TX)$ is said to be **Hamiltonian** if $W \lrcorner \omega$ is exact and is said to be **symplectic** if $W \lrcorner \omega$ is closed.

By definition all Hamiltonian vector fields are symplectic. However, the converse is not true.

Example 4.7. (Symplectic Vector Field that is not Hamiltonian)

The 2-torus \mathbb{T}^2 is a symplectic manifold when equipped with local 2-form $d\theta \wedge d\varphi$, where θ and φ are two different local coordinates for S^1 . The vector field $\frac{\partial}{\partial\theta}$ on \mathbb{T}^2 is symplectic but not Hamiltonian since $\frac{\partial}{\partial\theta} \lrcorner (d\theta \wedge d\varphi) = d\varphi$ is closed but not exact. The 1-form $d\varphi$ is locally exact, but not globally exact since φ is only defined on a proper open subset of S^1 . The same can be said about the vector field $\frac{\partial}{\partial\varphi}$.

Remark 4.8. The Lie bracket $[\cdot, \cdot]$ turns the subspace of Hamiltonian vector fields into a Lie algebra. In fact a stronger result holds; the Lie bracket of any two symplectic vector fields is Hamiltonian. Indeed, if X, Y are symplectic then, using the identity

$$[\mathcal{L}_X, \iota_Y] = \iota_{[X, Y]} = [\iota_X, \mathcal{L}_Y]$$

it follows that

$$\begin{aligned} [X, Y] \lrcorner \omega &= \iota_{[X, Y]} \omega \\ &= \mathcal{L}_X \iota_Y \omega - \iota_Y \mathcal{L}_X \omega \\ &= d(X \lrcorner Y \lrcorner \omega) + X \lrcorner d(Y \lrcorner \omega) - Y \lrcorner d(X \lrcorner \omega) \\ &= d(\omega(Y, X)) \end{aligned}$$

Just as the Euler-Lagrange equations determine the motions a Lagrangian system must satisfy, curves satisfying the Hamilton equations give the motions in a Hamiltonian system. We will see why this is below.

4.2 Hamilton's Equations

By Darboux's theorem we can find local coordinates $(q^1, \dots, q^n, p_1, \dots, p_n)$ in X such that $\omega = dq^j \wedge dp_j$. We have that $V_H = A^i \frac{\partial}{\partial q^i} + B_i \frac{\partial}{\partial p_i}$ for some $A^i, B_i \in C^\infty(X)$. It follows that $V_H \lrcorner \omega = A^i dp_i - B_i dq^i$. Since $dH = \frac{\partial H}{\partial q^i} dq^i + \frac{\partial H}{\partial p_i} dp_i$, it must be that $A^i = \frac{\partial H}{\partial p_i}$ and $B_i = -\frac{\partial H}{\partial q^i}$. Hence any integral curve $\gamma(t) = (\alpha(t), \beta(t))$ of V_H must satisfy

$$\frac{d}{dt} \alpha^i(t) = \left(\frac{\partial}{\partial p_i} H \right) (\gamma(t))$$

$$\frac{d}{dt} \beta_i(t) = - \left(\frac{\partial}{\partial q^i} H \right) (\gamma(t))$$

These are the **Hamilton equations**, a system of $2n$ first order ODE's. That is, a curve $\gamma : \mathbb{R} \rightarrow X$ is an integral curve for V_H if and only if $\gamma(t)$ satisfies Hamilton equations. We have shown that

$$V_H = \frac{\partial H}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H}{\partial q^i} \frac{\partial}{\partial p_i}$$

Definition 4.9. An integral curve $\gamma : \mathbb{R} \rightarrow X$ of V_H is called a **motion** of the Hamiltonian system (X, ω, H) . That is, γ is a motion if and only if γ satisfies Hamilton's equations.

A rather basic example of a Hamiltonian system is the simple pendulum. This was discussed in the Lagrangian setting as Example 3.25

Example 4.10. (The Simple Pendulum)

Consider the 1-dimensional manifold S^1 . Let (V, θ) be a chart in S^1 . We know that the cotangent bundle (T^*S^1, θ, ξ) is a symplectic manifold with symplectic 2-form $\omega = d\theta \wedge d\xi$. Consider the function $K : \mathbb{R} \rightarrow \mathbb{R}$ given by $\xi \mapsto \frac{\xi^2}{2ml^2}$ and $V : U \rightarrow \mathbb{R}$ given by $\theta \mapsto gml(1 - \cos \theta)$. Define the Hamiltonian to be

$$H : T^*S^1 \rightarrow \mathbb{R} \quad (\theta, \xi) \mapsto K(\xi) + U(\theta).$$

At first this definition of the Hamiltonian may seem ad hoc; however, after introducing the Legendre transform we will see where it comes from. In fact, the choice for naming the functions K and U above is to indicate that the Hamiltonian is to be thought of, in this setting, as the total energy. It follows that the Hamiltonian vector field is

$$V_H = \frac{\partial H}{\partial \xi} \frac{\partial}{\partial \theta} - \frac{\partial H}{\partial \theta} \frac{\partial}{\partial \xi} = mgl \sin \theta \frac{\partial}{\partial \theta} - \frac{\xi}{ml^2} \frac{\partial}{\partial \xi}$$

Hence if $\gamma(t) = (\alpha(t), \beta(t))$, where $\alpha(t) = \theta(\gamma(t))$ and $\beta(t) = \xi(\alpha(t))$, is a motion in this Hamiltonian system then $\gamma(t)$ satisfies the Hamilton equations

$$\frac{d}{dt}(\beta(t)) = \frac{\partial H}{\partial \xi} \Big|_{\gamma(t)} \implies \beta'(t) = mgl \sin(\alpha(t))$$

and

$$\frac{d}{dt}(\alpha(t)) = - \frac{\partial H}{\partial \theta} \Big|_{\gamma(t)} \implies \alpha'(t) = - \frac{\beta(t)}{ml^2}$$

Combining these two equations it follows that

$$\alpha''(t) = -\frac{g}{l} \sin \theta$$

which is precisely the motion of the pendulum as prescribed by Newton's second law. That is, the integral curves of the Hamiltonian vector field give the motions of this mechanical system.

4.3 The Poisson Bracket

Given a symplectic manifold (X, ω) , the Poisson bracket turns $C^\infty(X)$ into a Lie algebra. Suppose that the symplectic manifold also has a Hamiltonian $H \in C^\infty(X)$. In the same way the Lie bracket measures commutativity of vector fields, the Poisson bracket measures the commutativity of functions with the Hamiltonian vector field.

Definition 4.11. Given $f, g \in C^\infty(X)$ their **Poisson bracket** is defined to be

$$\{f, g\} := \omega(V_f, V_g).$$

By definition

$$\{f, g\} := \omega(V_f, V_g) = (V_f \lrcorner \omega)(V_g) = df(V_g) = V_g f.$$

Using this calculation, is not hard to verify that $\{\cdot, \cdot\}$ does indeed turn $C^\infty(X)$ into a Lie algebra. Moreover, there is a Leibniz rule;

$$\{f, gh\} = g\{f, h\} + \{f, g\}h.$$

A straightforward computation shows that $C^\infty(X) \ni H \mapsto V_H \in \Gamma(TX)$ is a Lie algebra anti-homomorphism.

Proposition 4.12. For $f \in C^\infty(X)$ we have $\{f, H\} = 0 \iff f \circ \theta_t = f$ on the domain of θ_t , where θ_t is the flow of V_H . That is, $\{f, H\} = 0$ if and only if f is constant along the integral curves of V_H .

Proof.

$$\begin{aligned} f \circ \theta_t = f &\iff \theta_t^* f = f \\ &\iff \mathcal{L}_{V_H} f = 0 \\ &\iff V_H f = 0 \\ &\iff (df)(V_H) = 0 \\ &\iff (V_f \lrcorner \omega)(V_H) = 0 \\ &\iff \omega(V_H, V_f) = 0 \\ &\iff \{H, f\} = 0 = \{f, H\} \end{aligned}$$

□

Definition 4.13. A function f such that $\{f, H\} = 0$ is called an **integral of motion**.

By definition, if two functions f_1, f_2 commute (with respect to $\{\cdot, \cdot\}$) then $\omega(X_{f_1}, X_{f_2}) = 0$, showing that given a collection of commuting integrals of motion, they generate an isotropic subspace of $T_p M$. But in the paragraph following Definition 2.6, we showed that given a subspace Y

of a vector space V we have $\dim V = \dim Y + \dim Y^\Omega$. If Y is isotropic then $Y \subset Y^\Omega$ so that an isotropic subspace has dimension at most half the dimension of M . A Hamiltonian system is called **(completely) integrable** if there exists n Poisson commuting independent integrals of motion $f_1 = H, f_2, \dots, f_n$.

Theorem 4.14. (Arnold-Lioville Theorem, [1])

Let (X, ω, H) be a completely integrable Hamiltonian system of dimension $2n$ with integrals of motion $f_1 = H, f_2, \dots, f_n$. Consider the function $f : TX \rightarrow \mathbb{R}^n$ defined by $f := (f_1, \dots, f_n)$. Let $c \in \mathbb{R}^n$ be a regular value of f . That is, c is a point in \mathbb{R}^n such that for every point in $f^{-1}(c)$ the differential of f is surjective. Then $f^{-1}(c)$ is a Lagrangian submanifold of X . Moreover we have

- (a) If the flows of the Hamiltonian vector fields X_{f_1}, \dots, X_{f_n} starting at a point $p \in f^{-1}(c)$ are complete, then the connected component of $f^{-1}(c)$ containing p is a homogeneous space for \mathbb{R}^n . This connected component has local coordinates $\varphi_1, \dots, \varphi_n$, called **angle coordinates**, for which the flows of X_{f_1}, \dots, X_{f_n} are linear.
- (b) There exists coordinates ψ_1, \dots, ψ_n , called **action coordinates**, such that each ψ_i is an integral of motion and also such that $\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_n$ form Darboux coordinates.

Proof. This is Theorem 18.12 in [1]. A proof of part (a) can be found in [1], page 110. For a proof of part (b) see [2], pages 271-274 and 279-281. \square

There is a rather simple expression of the Poisson bracket in Darboux coordinates.

Proposition 4.15. *Let (M, ω) be a symplectic manifold and let $(U, q^1, \dots, q^n, p_1, \dots, p_n)$ be a Darboux chart. Then for $f, g \in C^\infty(M)$ we have $\{f, g\} = \frac{\partial g}{\partial p_i} \frac{\partial f}{\partial q^i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i}$.*

Proof. In the above Darboux coordinates we have that $\omega = dq^i \wedge dp_i$. Hence

$$\begin{aligned} \{f, g\} &:= \omega(V_f, V_g) \\ &= (dq^i \wedge dp_i) \left(\frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial f}{\partial q^i} \frac{\partial}{\partial p_i}, \frac{\partial g}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial g}{\partial q^i} \frac{\partial}{\partial p_i} \right) \\ &= -\frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} - \left(-\frac{\partial g}{\partial p_i} \frac{\partial f}{\partial q^i} \right) \\ &= \frac{\partial g}{\partial p_i} \frac{\partial f}{\partial q^i} - \frac{\partial f}{\partial p_i} \frac{\partial g}{\partial q^i} \end{aligned}$$

\square

5 The Legendre Transform

We have seen the definition of motions in a Lagrangian system (M, L) . These are just curves whose derivatives are solutions to the Euler-Lagrange equations. We have also seen how motions are defined in a Hamiltonian system (X, ω, H) . These are just curves that satisfy the Hamilton equations. But given a Lagrangian system (M, L) we can always consider the cotangent bundle to get a symplectic manifold (T^*M, ω) . The Legendre transform provides the link between Lagrangian mechanics in (M, L) and Hamiltonian mechanics in $(T^*M, \omega = -d\alpha, H)$, where $H \in C^\infty(T^*M)$ will be defined below. Conversely, given a Hamiltonian system on a manifold which is a cotangent bundle, $(T^*M, \omega = -d\alpha, H)$ we can consider, under the Legendre transform, motions in an induced Lagrangian system (M, L) . In section 6 we will see how the Legendre transform relates the two statements of Noether's theorem.

5.1 The Legendre Transform on a Vector Space

Let V denote an n -dimensional vector space with ordered basis $\{e_1, \dots, e_n\}$ and let $\{v^1, \dots, v^n\}$ denote the coordinate functions. Fix $L \in C^\infty(V)$.

Definition 5.1. The Legendre transform associated to L is the map

$$\Phi_L : V \rightarrow V^* \quad p \mapsto \frac{\partial L}{\partial v}(p)$$

where $\frac{\partial L}{\partial v}(p)$ is the co-vector $(\frac{\partial L}{\partial v^1}(p), \dots, \frac{\partial L}{\partial v^n}(p)) \in T_p^*V \cong V^*$

In other words, $\Phi_L(p)$ is the Jacobian of L evaluated at p .

Definition 5.2. The dual function associated to L is the function $L^* : V^* \rightarrow \mathbb{R}$ defined by

$$L^* : V^* \rightarrow \mathbb{R} \quad \alpha \mapsto \sup\{\alpha \cdot p - L(p) \mid p \in V\}$$

Notice that if V^* has coordinates $\{\xi_1, \dots, \xi_n\}$ with respect to the dual basis $\{e^1, \dots, e^n\}$ then we can take the Legendre transform of L^* which is just $\Phi_{L^*}(p) = \text{Jac}(L^*)|_p = \left[\frac{\partial L^*}{\partial \xi_i}(p) \right]$.

Suppose now that $V = \mathbb{R}^n$. For the rest of this section, let (x^1, \dots, x^n) denote the standard coordinates on \mathbb{R}^n and let $(x^1, \dots, x^n, v^1, \dots, v^n)$ and $(x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ be the induced coordinates on $T\mathbb{R}^n$ and $T^*\mathbb{R}^n$ respectively. Fix a smooth function $L \in C^\infty(T\mathbb{R}^n) = C^\infty(\mathbb{R}^{2n})$. For what is to follow, let $x = (x^1, \dots, x^n)$, $v = (v^1, \dots, v^n)$, $a = (a_1, \dots, a_n)$ and $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ be arbitrary. For each $x \in \mathbb{R}^n$ the map L gives an induced map

$$L_x : \mathbb{R}^n \rightarrow \mathbb{R} \quad v \mapsto L(x, v)$$

Recall that the **Hessian** of L_x is the map $\text{Hess}(L_x) : \mathbb{R}^n \rightarrow L(\mathbb{R}^n, \mathbb{R}^n)$ defined by

$$\text{Hess}(L_x) := \left[\frac{\partial^2 L_x}{\partial v^i \partial v^j} \right]$$

Remark 5.3. For fixed $x \in \mathbb{R}^n$ the Legendre transform is a function from \mathbb{R}^n to \mathbb{R}^n

$$\Phi_{L_x} : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad p \mapsto \frac{\partial L_x}{\partial v}(p)$$

and so by definition of the Hessian it follows that

$$\text{Hess}(L_x) = \left[\frac{\partial^2 L_x}{\partial v^i \partial v^j} \right] = \text{Jac}(\Phi_{L_x})$$

Definition 5.4. A function $L \in C^\infty(T\mathbb{R}^n)$ is called **strongly convex** if for each $p \in \mathbb{R}^n$ the symmetric matrix $\text{Hess}(L_x(p))$ satisfies $u^T \text{Hess}(L_x(p))u > 0$ for all non-zero $u \in \mathbb{R}^n$ (i.e. $\text{Hess}(L_x(p))$ is a positive definite matrix).

Proposition 5.5. *If $L \in C^\infty(T\mathbb{R}^n)$ is strongly convex then L is strictly convex.*

Proof. Let $x \in \mathbb{R}^n$ be arbitrary. We need to show that $L_x : \mathbb{R}^n \rightarrow \mathbb{R}$ is strictly convex. For arbitrary, $p, q \in \mathbb{R}^n$ with $q \neq 0$ let $(L_x)_{p,q}$ denote the function

$$(L_x)_{p,q} : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto L_x(p + tq).$$

Notice that L_x is strictly convex if and only if $(L_x)_{p,q}$ is strictly convex for all $p, q \in \mathbb{R}^n$. But a standard calculation shows that $(q^T \cdot \text{Hess}(L_x)_{p,q} \cdot q) = (L_x)''_{p,q}$. By assumption, it follows that $(L_x)''_{p,q}(t) > 0$ for all $t \in \mathbb{R}$ and $p, q \in \mathbb{R}^n$. Thus, from basic calculus, it follows that $(L_x)_{p,q}$ is strictly convex for all $p, q \in \mathbb{R}^n$. \square

Proposition 5.6. *Fix $x \in \mathbb{R}^n$ and suppose that $L_x : \mathbb{R}^n \rightarrow \mathbb{R}$ is strongly convex. Then the following are equivalent*

1. L_x has a critical point (i.e. there exists $v_0 \in \mathbb{R}^n$ such that $\frac{\partial L_x}{\partial v^i}(v_0) = 0$ for all $i = 1, \dots, n$).
2. L_x has a local minimum
3. L_x has a unique global minimum

Proof. (1) \implies (2) Suppose that v_0 is a critical point L_x . By hypothesis $\text{Hess}(L_x)|_{v_0}$ is positive definite and so has only positive eigenvalues. Thus by the second derivative test (see [14], Theorem 6.37) L has a local minimum at p .

(2) \implies (3) Suppose that v_0 is a local minimum of L . Then by definition there exists a neighbourhood $U \subset \mathbb{R}^n$ such that $v_0 \in U$ and $L(v_0) \leq L(u)$ for all $u \in U$. Suppose that v_0 is not a global minimum. Then there exists $w \in \mathbb{R}^n$ such that $L(w) < L(v_0)$. But then for arbitrary $\theta \in (0, 1)$, by

Proposition 5.5, we have that $L((1 - \theta)v_0 + \theta w) < L(v_0) - \theta L(v_0) + \theta L(w) < L(v_0)$. But this is a contradiction since we can choose θ sufficiently small so that $(1 - \theta)v_0 + \theta w \in U$.

(3) \implies (1) This is known from basic calculus. \square

Proposition 5.7. *Fix an arbitrary $x \in \mathbb{R}^n$. If $L_x \in C^\infty(T_x \mathbb{R}^n) = C^\infty(\mathbb{R}^n)$ is strongly convex then $\Phi_{L_x} : T_x \mathbb{R}^n \rightarrow \Phi_{L_x}(T_x \mathbb{R}^n)$ is a diffeomorphism.*

Proof. By definition, for arbitrary $v_0 \in T_x \mathbb{R}^n \cong \mathbb{R}^n$ we have $\Phi_{L_x}(v_0) = \frac{\partial L_x}{\partial v}(v_0) \in T_x^* \mathbb{R}^n = \mathbb{R}^n$. By assumption, the Jacobian of Φ_{L_x} is positive definite. That is, $\text{Hess}(L_x)|_{v_0}$ is positive definite so in particular

$$\det \left[\frac{\partial}{\partial v^j} \Phi_{L_x}(v_0) \right] = \det \left[\frac{\partial}{\partial v^j} \frac{\partial}{\partial v^i} L_x(v_0) \right] > 0$$

Thus, by the inverse function theorem, Φ_{L_x} is a local diffeomorphism. Since a bijective local diffeomorphism is a diffeomorphism, it suffices to show that $\Phi_{L_x} : T_x \mathbb{R}^n = \mathbb{R}^n \rightarrow \Phi_{L_x}(\mathbb{R}^n)$ is injective. So suppose that $p, q \in \mathbb{R}^n = T_x \mathbb{R}^n$ are such that $p \neq q$. Let $w = q - p$ so that $w \neq 0$. Since L_x is smooth, we have that L_x is smooth on the line segment $\{p + tw ; 0 \leq t \leq 1\}$. By the chain rule

$$\frac{d}{dt} \Phi_{L_x}(p + tw) = (\text{Jac}(\Phi_{L_x})(p + tw)) w = \text{Hess}(L_x(p + tw))w$$

We also have that $\Phi_{L_x}(q) - \Phi_{L_x}(p) = \int_0^1 \frac{d}{dt} \Phi_{L_x}(p + tw) dt$. Putting this together yields

$$\begin{aligned} w^T (\Phi_{L_x}(q) - \Phi_{L_x}(p)) &= w^T \left(\int_0^1 \frac{d}{dt} \Phi_{L_x}(p + tw) dt \right) \\ &= w^T \left(\int_0^1 \text{Hess}(L_x(p + tw)) w dt \right) \\ &= \int_0^1 w^T \text{Hess}(L_x(p + tw)) w dt \end{aligned}$$

However, $\text{Hess}(L_x)$ is positive definite on \mathbb{R}^n and so this last expression is positive. Thus it can't be that $\Phi_{L_x}(p) = \Phi_{L_x}(q)$. Hence $\Phi_{L_x} = \frac{\partial}{\partial v} L_x$ is injective. \square

Proposition 5.8. *If L_x is strongly convex then for all $\xi \in \Phi_{L_x}(\mathbb{R}^n)$ we have*

$$L_x^*(\xi) = \xi \cdot \Phi_{L_x}^{-1}(\xi) - L_x(\Phi_{L_x}^{-1}(\xi))$$

Proof. Fix $\xi \in \Phi_{L_x}(\mathbb{R}^n) \subset T_x^* \mathbb{R}^n = \mathbb{R}^n$. Consider the function $g : \mathbb{R}^n \rightarrow \mathbb{R}$ defined by $g(v) = \xi \cdot v - L_x(v)$. It's clear that g is smooth. We have that

$$\frac{\partial}{\partial v} g(v) = \xi - \frac{\partial}{\partial v} L_x(v) = \xi - \Phi_{L_x}(v)$$

and so

$$\text{Hess}(g)|_v = -\text{Hess}(L_x)|_v$$

Hence, by hypothesis, it follows $\text{Hess}(g(v))$ is negative definite. Also, by Proposition 5.7 there exists a unique $q \in \mathbb{R}^n$ such that $\Phi_{L_x}(q) = \frac{\partial L_x}{\partial v}(q) = \xi$. That is, $\frac{\partial}{\partial v}g(q) = 0$. But since $\text{Hess}(g)$ is negative definite, by Proposition 5.6, it follows q is a global maximum for g . Hence $g(a) \leq g(q)$ for all $a \in \mathbb{R}^n$. But then

$$\begin{aligned} L_x^*(\xi) &= \sup\{\xi a - f(a) \mid a \in \mathbb{R}^n\} \\ &= \sup\{g(a) \mid a \in \mathbb{R}^n\} \\ &= g(q) \\ &= \xi(\Phi_{L_x}^{-1}(\xi)) - L_x(\Phi_{L_x}^{-1}(\xi)) \end{aligned}$$

□

Theorem 5.9. *If L_x is strongly convex then $\Phi_{L_x^*} = \Phi_{L_x}^{-1}$.*

Proof. By definition, the Legendre transform of L^* is

$$\Phi_{L_x^*} : T_x^*\mathbb{R}^n \rightarrow T_x\mathbb{R}^n, \quad \xi \mapsto \frac{\partial L_x^*}{\partial \xi}$$

Proposition 5.8 showed that for all $\xi \in \Phi_{L_x}(U)$

$$L_x^*(\xi) = \xi \cdot \Phi_{L_x}^{-1}(\xi) - L_x(\Phi_{L_x}^{-1}(\xi))$$

Thus

$$\begin{aligned} \Phi_{L_x^*}(\xi) &= \frac{\partial L_x^*}{\partial \xi}(\xi) \\ &= \Phi_{L_x}^{-1}(\xi) + \xi \left(\frac{\partial}{\partial \xi} \Phi_{L_x}^{-1}(\xi) \right) - \left(\frac{\partial L_x}{\partial v}(\Phi_{L_x}^{-1}(\xi)) \cdot \left(\frac{\partial \Phi_{L_x}^{-1}}{\partial \xi}(\xi) \right) \right) && \text{by the chain rule} \\ &= \Phi_{L_x}^{-1}(\xi) + \xi \left(\frac{\partial}{\partial \xi} \Phi_{L_x}^{-1}(\xi) \right) - \left(\Phi_{L_x}(\Phi_{L_x}^{-1}(\xi)) \cdot \left(\frac{\partial \Phi_{L_x}^{-1}}{\partial \xi}(\xi) \right) \right) && \text{by definition} \\ &= \Phi_{L_x}^{-1}(\xi) + \xi \left(\frac{\partial}{\partial \xi} \Phi_{L_x}^{-1}(\xi) \right) - \xi \cdot \left(\frac{\partial \Phi_{L_x}^{-1}}{\partial \xi}(\xi) \right) \\ &= \Phi_{L_x}^{-1}(\xi) \end{aligned}$$

□

Remark 5.10. Let $L \in C^\infty(\mathbb{R}^n)$ be a strongly convex function. We say that L has **quadratic growth at infinity** if there exists a positive definite quadratic form Q on \mathbb{R}^n and a constant K such that $L(p) \geq Q(p) - K$ for all $p \in \mathbb{R}^n$. If L has quadratic growth at infinity, then $\Phi_L(\mathbb{R}^n) = (\mathbb{R}^n)^*$. That is, if L is strongly convex and has quadratic growth at infinity, then $\Phi_L : T\mathbb{R}^n \rightarrow T^*\mathbb{R}^n$ is a diffeomorphism. This is exercise 54 in [1]. We do not give a proof of this result as we do not need it for what is to follow.

Using the above theorems we can show that the dual function of a strongly convex Lagrangian is strongly convex.

Proposition 5.11. *If $L \in C^\infty(T\mathbb{R}^n)$ is strongly convex, then $H = L^* \in C^\infty(T^*\mathbb{R}^n)$ is strongly convex.*

Proof. Suppose that $L \in C^\infty(\mathbb{R}^n)$ is strongly convex. We need to show that $H = L^* \in C^\infty((\mathbb{R}^n)^*)$ is strongly convex. That is, for fixed $x \in \mathbb{R}^n$, we need to show that for all $\xi \in \mathbb{R}^n$ we have $\text{Hess}(H_x(\xi))$ is positive definite. We showed in the proof of Theorem 5.9 that $\frac{\partial L_x^*}{\partial \xi}(\xi) = \Phi_{L_x}^{-1}(\xi)$. We also have that

$$\Phi_{L_x}(\Phi_{L_x}^{-1}(\xi)) = \Phi_{L_x}(\Phi_{H_x}(\xi)) = \xi.$$

Differentiating this equality with respect to ξ , the chain rule gives that

$$\left(\frac{\partial \Phi_{L_x}}{\partial v}(\Phi_{H_x}(\xi)) \right) \left(\frac{\partial}{\partial \xi} \Phi_{H_x}(\xi) \right) = 1.$$

Since $\Phi_{L_x} = \frac{\partial L_x}{\partial v}$ and $\Phi_{H_x} = \frac{\partial L_x^*}{\partial \xi} = \frac{\partial H_x}{\partial \xi}$ this equation is the same as

$$\left(\frac{\partial^2 L_x}{(\partial v)^2}(\Phi_{H_x}(\xi)) \right) \left(\frac{\partial^2 H_x}{(\partial \xi)^2}(\xi) \right) = 1.$$

That is we have shown that

$$\text{Hess}(H_x(\xi)) = \text{Hess}(L_x(\Phi_{H_x}(\xi)))^{-1}.$$

Since the inverse of a positive definite matrix is positive definite, the result follows. \square

Corollary 5.12. (Involutivity of the Legendre Transform and Dual Function)

Let $L \in C^\infty(\mathbb{R}^n)$ be strongly convex. If $H = L^*$, then $H^* = L$. In particular, this means that $\Phi_{H^*} = \Phi_{L^{**}} = \Phi_L$.

Proof. Let $L \in C^\infty(\mathbb{R}^n)$ be strongly convex. From Proposition 5.11 we have that $H = L^*$ is strongly convex. Hence

$$\begin{aligned} H^*(x, v) &= v \cdot \Phi_{H_x}^{-1}(v) - H_x(\Phi_{H_x}^{-1}(v)) && \text{by Proposition 5.8 and 5.11} \\ &= v \cdot \Phi_{H_x}^{-1}(v) - ((\Phi_{H_x}^{-1}(v)) \cdot (\Phi_{L_x}^{-1}(\Phi_{H_x}^{-1}(v)))) - L_x(\Phi_{L_x}^{-1}(\Phi_{H_x}^{-1}(v))) && \text{by Proposition 5.8} \\ &= v \cdot \Phi_{H_x}^{-1}(v) - ((\Phi_{H_x}^{-1}(v)) \cdot v - L_x(v)) && \text{since } \Phi_{L_x}^{-1} = \Phi_{H_x} \\ &= L(x, v) \end{aligned}$$

\square

5.2 The Legendre Transform on Manifolds

The Legendre transform can be extended naturally to act on manifolds since, at every point, the tangent and cotangent spaces are vector spaces.

Let M be a n -dimensional manifold and (U, x^1, \dots, x^n) an arbitrary coordinate chart. By definition, U is diffeomorphic to an open subset of \mathbb{R}^n . We have the induced coordinate charts $(TU, x^1, \dots, x^n, v^1, \dots, v^n)$ and $(T^*U, x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ on TM and T^*M respectively, and we know that $TU \cong U \times \mathbb{R}^n \cong T^*U$. Suppose that $L \in C^\infty(TU)$ is strongly convex. We define the Legendre transform associated to L to be the map

$$\Phi_L : TU \rightarrow \Phi_L(TU) \quad (x, v) \mapsto \frac{\partial L}{\partial v}(x, v).$$

For a fixed $x \in M$, we have that $L_x \in C^\infty(T_x U)$ is strongly convex with respect to v^1, \dots, v^n . The Legendre transform induces the map

$$\Phi_{L_x} : T_x U \rightarrow T_x^* U \quad W_x \mapsto \Phi_{L_x}(W_x) = \frac{\partial L_x}{\partial v}(W_x)$$

That is, $\frac{\partial L_x}{\partial v}(W_x)$ is the n -tuple $(\frac{\partial L_x}{\partial v^1}(W_x), \dots, \frac{\partial L_x}{\partial v^n}(W_x))$. For each $x \in M$ the dual function associated to L is again defined to be the map

$$L_x^* : T_x^* U \rightarrow \mathbb{R} \quad \xi_x \mapsto \sup\{\xi_x \cdot W_x - L_x(W_x) ; W_x \in T_x U\}$$

All of the results from the previous section still hold. That is, for each $x \in M$, we have

- L_x has a critical point $\iff L_x$ has a local minimum $\iff L_x$ has a unique global minimum.
- $\Phi_{L_x} : T_x U \rightarrow \Phi_{L_x}(T_x U)$ is a diffeomorphism.
- For all $\xi_x \in \Phi_{L_x}(T_x U)$ we have $L_x^*(\xi_x) = \xi_x \cdot \Phi_{L_x}^{-1}(\xi_x) - L_x(\Phi_{L_x}^{-1}(\xi_x))$.
- $\Phi_{L_x^*} = \Phi_{L_x}^{-1}$.
- If $L \in C^\infty(TU)$ is strongly convex then $H = L^* \in C^\infty(\Phi_L(TU))$ is strongly convex.
- If $L \in C^\infty(TU)$ is strongly convex and $H = L^* \in C^\infty(\Phi_L(TU))$ then $H^* = L$. That is, the dual function and Legendre transform are involutive.

5.3 The Legendre Transform Relates Lagrangian and Hamiltonian Mechanics

Let (M, L) be an arbitrary Lagrangian system, where $L \in C^\infty(TM)$ is strongly convex. In the previous section we defined $L^* \in C^\infty(T^*M)$, the dual function of L . Moreover, we know that T^*M is a symplectic manifold when equipped with the canonical 2-form ω . Hence, we see that the Hamiltonian system $(T^*M, \omega, H = L^*)$ arises naturally from the Lagrangian system (M, L) . Similarly, given a Hamiltonian system of the form (T^*M, ω, H) for some strongly convex $H \in$

$C^\infty(T^*M)$, we can define the dual function $H^* \in C^\infty(TM)$. This gives the Lagrangian system $(M, L = H^*)$. Since the Legendre transform and the dual function are involutive, we see that these induced systems are well defined and ‘inverse’ to each other. This motivates the following definition.

Definition 5.13. Given a Lagrangian system (M, L) the **induced Hamiltonian system** is the triple $(T^*M, \omega, H := L^*)$ where ω is the canonical 2-form and T^*M . Similarly, given a Hamiltonian system of the form (T^*M, ω, H) , the **induced Lagrangian system** is the pair $(M, L := H^*)$.

Remark 5.14. Let (U, x^1, \dots, x^n) be a coordinate chart in a manifold M . We have the induced coordinate charts $(TU, x^1, \dots, x^n, v^1, \dots, v^n)$ and $(T^*U, x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ on TM and T^*M respectively. Proposition 5.7 showed that $\Phi_L : TU \rightarrow \Phi_L(TU)$ is a diffeomorphism, while Theorem 5.9 showed that $\Phi_L^{-1} = \Phi_{L^*}$. Hence, in the induced Hamiltonian system $(T^*M, \omega, H = L^*)$ we have the coordinate chart $(\Phi_L(TU), x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ where each ξ_i satisfies $\xi_i = \frac{\partial L}{\partial v^i}$. We also have that $\Phi_L^{-1} = \Phi_H$. Similarly, if we are given a Hamiltonian system of the form (T^*M, ω, H) , then the Legendre transform $\Phi_H : T^*U \rightarrow \Phi_H(T^*U)$ gives an induced coordinate chart $(\Phi_H(T^*U), x^1, \dots, x^n, v^1, \dots, v^n)$ where each $v^i = \frac{\partial H}{\partial \xi_i}$. Also, we have that $\Phi_H^{-1} = \Phi_L$.

Given a natural Lagrangian system, the Hamiltonian function in the induced Hamiltonian system is always the total energy.

Proposition 5.15. Let (M, g) be a Riemannian manifold with Lagrangian $L = K - U \in C^\infty(T^*M)$. Then $H := L^* = E = K + U$.

Proof. By definition, $L = \frac{1}{2}mg_{ij}v^i v^j - U$. For $(x, v) \in TM$, let $(x, \xi) = \Phi_L(x, v) \in \Gamma(T^*M)$. That is,

$$\begin{aligned} \xi_k &= \frac{\partial L}{\partial v^k} \\ &= \frac{\partial}{\partial v^k} \left(\frac{1}{2}mg_{ij}v^i v^j \right) && \text{since } U \text{ is independent of } v^1, \dots, v^n \\ &= \frac{1}{2}mv^j g_{ij} \delta_k^i + \frac{1}{2}mv^i g_{ij} \delta_k^j && \text{by the product rule} \\ &= mg_{ik}v^i \end{aligned}$$

Letting $g^{ij} = [g^{-1}]_{ij}$ it follows that

$$v^i = \frac{1}{m}g^{ik}\xi_k.$$

Since $\Phi_L^{-1}(x, \xi) = \Phi_L^{-1}(\Phi_L(x, v)) = (x, v)$, by Proposition 5.8 it follows that

$$\begin{aligned} H(x, \xi) &= L^*(x, \xi) \\ &= \xi_i v^i - L(x, v) \end{aligned}$$

$$\begin{aligned}
&= mg_{ij}v^jv^i - \frac{1}{2}mg_{ij}v^iv^j + U(x) \\
&= \frac{1}{2}mg_{ij}v^iv^j + U(x) \\
&= K + U
\end{aligned}$$

□

Remark 5.16. Notice that if the metric is the standard one and the manifold is \mathbb{R}^n , then in $TM = \mathbb{R}^{2n}$ motions are described by their position and velocity coordinates. However, in the induced Hamiltonian system we have $\xi_i = mv^i$ and so motions in here are described at each time by specifying position and momentum coordinates.

A simple calculation, which we show now, demonstrates that if the natural Lagrangian is time independent then the total energy is conserved. In section 6 we will see that time independence of the natural Lagrangian can be thought of as a ‘symmetry’ and so conservation of energy can also be seen as a consequence of Noether’s theorem.

Proposition 5.17. *In a natural Lagrangian system, if the Lagrangian is independent of time then energy is conserved. That is, if $L(x^1, \dots, x^n, v^1, \dots, v^n, t)$ is such that $\frac{\partial L}{\partial t} = 0$ then $\frac{d}{dt}H = 0$.*

Proof. Let $\gamma(t)$ be a motion in (M, L) . That is, suppose $\tilde{\gamma}(t) = (\gamma(t), \gamma'(t))$ satisfies the Euler Lagrange equations. We have that

$$\begin{aligned}
\frac{d}{dt}L(\gamma(t), \gamma'(t), t) &= \frac{\partial L}{\partial x^i} \Big|_{(\tilde{\gamma}(t), t)} \gamma'(t) + \frac{\partial L}{\partial v^i} \Big|_{(\tilde{\gamma}(t), t)} \gamma''(t) + \frac{\partial L}{\partial t} \Big|_{(\tilde{\gamma}(t), t)} \\
&= \frac{d}{dt} \frac{\partial L}{\partial v^i} \Big|_{(\tilde{\gamma}(t), t)} \gamma'(t) + \frac{\partial L}{\partial v^i} \Big|_{(\tilde{\gamma}(t), t)} \gamma''(t) + \frac{\partial L}{\partial t} \Big|_{(\tilde{\gamma}(t), t)} \quad \text{by the Euler-Lagrange equations} \\
&= \frac{d}{dt} \left(\frac{\partial L}{\partial v^i} \Big|_{(\tilde{\gamma}(t), t)} \gamma'(t) \right) + \frac{\partial L}{\partial t} \Big|_{(\tilde{\gamma}(t), t)} \quad \text{by the product rule} \\
&= \frac{d}{dt} (\xi_i(\tilde{\gamma}(t))v^i(\tilde{\gamma}(t))) + \frac{\partial L}{\partial t} \Big|_{(\tilde{\gamma}(t), t)} \quad \text{by definition}
\end{aligned}$$

That is

$$\frac{d}{dt} (L(x, v, t) - \xi_i v^i) = \frac{\partial L}{\partial t} \Big|_{(\tilde{\gamma}(t), t)}.$$

However, by proposition 5.8 we have that $L(x, v, t) - \xi_i v^i = -L^* := -H$. Hence if $\frac{\partial L}{\partial t} = 0$ then $\frac{d}{dt}H = 0$. By Proposition 5.15 this means that the total energy is conserved. □

The first example we give that demonstrates how the Legendre transform relates Hamiltonian mechanics and Lagrangian mechanics is by showing how it translates motions in one formulation to motions in the other. To see this we first need the following lemma.

Lemma 5.18. Assume that $L \in C^\infty(TM)$ is strongly convex. As above, if $(TU, x^1, \dots, x^n, v^1, \dots, v^n)$ is a coordinate chart on TM we get the induced chart $(\Phi_L(TU), x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ on T^*U , where by definition $(x, \xi) = \Phi_L(x, v)$. Let $H = L^* \in C^\infty(T^*U)$. The claim is that $\frac{\partial L}{\partial x}(x, v) = -\frac{\partial H}{\partial \xi}(x, \xi)$

Proof. By Proposition 5.8, $H_x(\xi) = L_x^*(\xi) = \xi(\Phi_{L_x}^{-1}(\xi)) - L(\Phi_{L_x}^{-1}(\xi))$. However, $\Phi_{L_x}^{-1}(\xi) = \Phi_{L_x}^{-1}(\xi) = \Phi_{L_x}^{-1}(\Phi_{L_x}(v)) = v$ and so

$$H(x, \xi) = \xi \cdot v - L(x, v) = \xi_i v^i - L(x, v) \quad (5.1)$$

We know that $\frac{\partial v^j}{\partial x^i} = 0$ for all $1 \leq i, j \leq n$; however, by definition ξ is dependent on x and v . Hence, taking the total derivative of $H(x, \xi)$ with respect to x^i , the left hand side of (5.1) is

$$\frac{\partial H}{\partial x^i}(x, \xi) + \frac{\partial H}{\partial \xi_j}(x, \xi) \frac{\partial \xi_j}{\partial x^i}(x, v)$$

while the total derivative of the right hand side is

$$\frac{\partial \xi_j}{\partial x^i}(x, v) v^j - \left(\frac{\partial L}{\partial x^i}(x, v) \right)$$

However, by hypothesis we have that $\frac{\partial H}{\partial \xi_i}(x, \xi) = (\Phi_L^{-1}(\Phi_L(x, v)))^i = v^i$ and so combining these equalities finishes the proof. \square

This result gives us the following two theorems.

Theorem 5.19. If a curve $\gamma : \mathbb{R} \rightarrow U$ satisfies the Euler-Lagrange equations on some chart $U \subset M$, then $\Phi_L \circ \tilde{\gamma} : [a, b] \rightarrow T^*M$ is an integral curve of the Hamiltonian vector field V_H .

Proof. Let $(U, x_1, \dots, x_n) \subset M$ be an arbitrary chart. We have the induced charts $(TU, x^1, \dots, x^n, v^1, \dots, v^n)$ and $(\Phi_L(TU), x^1, \dots, x^n, \xi_1, \dots, \xi_n)$ on TM and T^*M respectively. By hypothesis, $\tilde{\gamma}(t) = (\gamma(t), \gamma'(t))$ satisfies the Euler-Lagrange equations. That is

$$\frac{\partial L}{\partial x^i}(\tilde{\gamma}(t)) = \frac{d}{dt} \left(\frac{\partial L}{\partial v^i}(\tilde{\gamma}(t)) \right)$$

Let $\Psi(t) = \Phi_L(\tilde{\gamma}(t)) = \Phi_L(\gamma(t), \gamma'(t)) = (\gamma(t), \Phi_{L_{\gamma(t)}}(\gamma'(t)))$. It needs to be shown that $\Psi(t)$ satisfies Hamilton's equations. That is, it needs to be shown that

$$\frac{d}{dt} \gamma^i(t) = \frac{\partial H}{\partial \xi_i}(\Psi(t))$$

$$\frac{d}{dt} (\Phi_{L_{\gamma(t)}}(\gamma'(t)))_i(t) = -\frac{\partial H}{\partial x^i}(\Psi(t))$$

But since $\Psi(t) = \Phi_L(\tilde{\gamma}(t))$ it follows that $\Phi_H(\Psi(t)) = \tilde{\gamma}(t)$. Hence

$$\gamma'(t) = \frac{\partial H}{\partial \xi}(\Psi(t))$$

This is precisely the first line of Hamilton's equations. The second line also holds since

$$\begin{aligned}
\frac{d}{dt}(\Phi_{L_{\gamma(t)}}(\gamma'(t)))_i &= \frac{d}{dt} \frac{\partial L}{\partial v^i}(\tilde{\gamma}(t)) && \text{by definition} \\
&= \frac{\partial L}{\partial x^i}(\tilde{\gamma}(t)) && \text{by the Euler-Lagrange equations} \\
&= -\frac{\partial H}{\partial x^i}(\gamma(t), \Phi_{L_{\gamma(t)}}(\gamma'(t))) && \text{by the Lemma} \\
&= -\frac{\partial H}{\partial x^i}(\Psi(t)) && \text{by definition}
\end{aligned}$$

□

A stronger version of the converse is also true:

Theorem 5.20. *Given a Lagrangian system (M, L) , where $L \in C^\infty(TM)$ is strongly convex, let $(T^*M, \omega, H = L^*)$ be the induced Hamiltonian system. If $\Psi : \mathbb{R} \rightarrow T^*M$ is an integral curve for V_H then $\Psi = \Phi_L \circ \gamma$ for some motion γ in (M, L) .*

Proof. Let $\Psi(t) = (\alpha(t), \beta(t)) \in T^*M$ be an integral curve for V_H . Then $\Psi(t)$ satisfies the Hamilton equations

$$\begin{aligned}
\frac{d}{dt}\alpha^i(t) &= \frac{\partial H}{\partial \xi_i}(\Psi(t)) \\
\frac{d}{dt}\beta_i(t) &= -\frac{\partial H}{\partial x^i}(\Psi(t))
\end{aligned}$$

Here α is a curve $\alpha : \mathbb{R} \rightarrow M$. It follows that

$$\tilde{\alpha}(t) = (\alpha(t), \alpha'(t)) = (\alpha(t), \frac{\partial H}{\partial \xi}(\Psi(t))) = (\alpha(t), \Phi_{H_{\alpha(t)}}(\beta(t)))$$

so that

$$\Phi_L(\tilde{\alpha}(t)) = (\alpha(t), \Phi_{L_{\alpha(t)}}\Phi_{H_{\alpha(t)}}(\beta(t)) = (\alpha(t), \beta(t)) = \Psi(t)$$

It suffices to show that $\tilde{\alpha}(t)$ satisfies the Euler-Lagrange equations. Indeed

$$\begin{aligned}
\frac{\partial L}{\partial x}(\tilde{\alpha}(t)) &= \frac{\partial L_{\alpha(t)}}{\partial x}(\frac{d}{dt}\alpha(t)) && \text{by definition} \\
&= -\frac{\partial H_{\alpha(t)}}{\partial x}(\beta(t)) && \text{by the Lemma 5.17} \\
&= -\frac{\partial H}{\partial x}(\Psi(t)) && \text{by definition} \\
&= \frac{d}{dt}(\beta(t)) && \text{by Hamilton's equations} \\
&= \frac{d}{dt}\Phi_{L_{\alpha(t)}}\left(\Phi_{H_{\alpha(t)}}(\beta(t))\right) && \text{since } \Phi_{L^*} = \Phi_L^{-1} \\
&= \frac{d}{dt}\Phi_{L_{\alpha(t)}}\left(\frac{\partial H}{\partial \xi}(\Psi(t))\right) && \text{by definition} \\
&= \frac{d}{dt}\Phi_{L_{\alpha(t)}}\left(\frac{d}{dt}(\alpha(t))\right) && \text{by Hamilton's equation}
\end{aligned}$$

$$= \frac{d}{dt} \frac{\partial L}{\partial v}(\tilde{\alpha}(t)) \quad \text{by definition}$$

□

Example 5.21. (Geodesic Flow in Hamiltonian Mechanics)

We can now show how the Legendre transform relates the concept of geodesic flow in Lagrangian and Hamiltonian mechanics. Let (M, g, L) be a natural Lagrangian system, where $L \in C^\infty(TM)$ is strongly convex. Recall that in section 2.7 we derived the geodesic flow as the symplectomorphism generated by the Riemann distance function. In other words, we set our Lagrangian to be

$$L : TM \rightarrow \mathbb{R} \quad (x, V_x) \mapsto \frac{1}{2}g_x(V_x, V_x).$$

As demonstrated in Example 3.19 if the net force on the mechanical system is 0, then the solutions to the Euler-Lagrange equations are geodesics. Consider what happens if we translate this system into the Hamiltonian setting. Using the argument in the proof of Proposition 5.15 we have that

$$L^*(x, \xi) = \frac{1}{2}g^{ij}\xi_i\xi_j.$$

By definition our Hamiltonian vector field is

$$V_H := \frac{\partial H}{\partial \xi_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial \xi_i}.$$

The integral curves, $\gamma(t) = (x(t), \xi(t))$, of V_H must satisfy Hamilton's equations:

$$\frac{d}{dt}x^k(t) = \frac{\partial H}{\partial \xi_k}$$

$$\frac{d}{dt}\xi_k(t) = -\frac{\partial H}{\partial x^k}$$

We have that

$$\frac{\partial H}{\partial \xi_k} = \frac{1}{2}g^{ij}\delta_i^k\xi_j + \frac{1}{2}g^{ij}\xi_i\delta_j^k = g^{kj}\xi_j \quad (5.2)$$

and

$$-\frac{\partial H}{\partial x^k} = -\frac{1}{2}\frac{\partial g^{ij}}{\partial x^k}\xi_i\xi_j \quad (5.3)$$

To make the notation clearer, we will denote the time derivative using dot notation. If $\gamma(t)$ satisfies Hamilton's equations then by equation (5.2) we have that $\xi_k = g_{ak}\dot{x}^a$. Plugging this into the second line of Hamilton's equations and using (5.3) we get that

$$\dot{\xi}_k = \frac{\partial g_{ak}}{\partial x^q}\dot{x}^a\dot{x}^q + g_{ak}\ddot{x}^a = -\frac{1}{2}\frac{\partial g^{ij}}{\partial x^k}g_{ia}g_{jp}\dot{x}^a\dot{x}^p \quad (5.4)$$

We can simplify this expression using the following claim.

Claim 5.22. We have that $-\frac{\partial g^{ij}}{\partial x^k} g_{ia} g_{jp} = \frac{\partial g_{ap}}{\partial x^k}$

Proof. We know that $g_{ap} g^{pj} = \delta_a^j$. Differentiating this with respect to x^k we get that

$$\frac{\partial g_{ap}}{\partial x^k} g^{pj} + g_{ap} \frac{\partial g^{pj}}{\partial x^k} = 0.$$

Multiplying both sides by g_{ji} and summing over j gives

$$\frac{\partial g_{ap}}{\partial x^k} \delta_i^p = -\frac{\partial g^{pj}}{\partial x^k} g_{ap} g_{ij}$$

□

Using this claim, equation (5.4) becomes

$$\frac{\partial g_{ak}}{\partial x^q} \dot{x}^a \dot{x}^q + g_{ak} \ddot{x}^a = \frac{1}{2} \frac{\partial g_{ap}}{\partial x^k} \dot{x}^a \dot{x}^p.$$

Rearranging, we get that if $\gamma(t)$ satisfies Hamilton's equations then

$$\ddot{x}^b = g^{kb} \left(\frac{1}{2} \frac{\partial g_{ap}}{\partial x^k} \dot{x}^a \dot{x}^p - \frac{\partial g_{ak}}{\partial x^q} \dot{x}^a \dot{x}^q \right) = -\frac{1}{2} g^{kb} \left(\frac{\partial g_{ak}}{\partial x^p} \dot{x}^a \dot{x}^p + \frac{\partial g_{pk}}{\partial x^a} \dot{x}^a \dot{x}^p - \frac{\partial g_{ap}}{\partial x^k} \dot{x}^a \dot{x}^p \right)$$

This is precisely the geodesic equation. Hence, a curve satisfying the Hamilton equations is a geodesic. Conversely, let $\gamma(t) = (x^1(t), \dots, x^n(t))$ be a geodesic in M . Then if we set $\xi_k = g_{ak} \dot{x}^a$, applying the above argument to $\psi(t) = (x^1(t), \dots, x^n(t), \xi_1(t), \dots, \xi_n(t))$ shows that ψ is an integral curve of V_H .

Remark 5.23. As in the previous example, consider a Lagrangian of the form $L = K$. Then, as in the proof of Proposition 5.15, we have for fixed $x \in M$

$$(\Phi_L(x, v))_i = g_{ji} v^j$$

That is, for $W_x = W^i \frac{\partial}{\partial x^i} \big|_x \in T_x M$ we have that

$$\Phi_{L_x}(W_x) = g_{ji} W^i dx^j.$$

Hence in this case the Legendre transform is just the musical isomorphism

$$\Phi_{L_x} : T_x M \rightarrow T_x^* M \quad W \mapsto g(W, \cdot).$$

In section 2.7 we showed that the symplectomorphism generated by the Riemann distance function was the geodesic flow

$$\varphi : TM \rightarrow TM \quad V \mapsto (\gamma_V(1), \gamma_V'(1)).$$

But in order to find this symplectomorphism we identified T^*M with TM via the musical isomorphism. In fact, now we can see that all we were doing in that section was solving the Hamilton equations. We were trying to find V and W such that $g(V, \cdot) = d_a L$ and $g(W, \cdot) = -d_b L$. We took geodesics (motions in (M, L)) and mapped them under the Legendre transform to motions in (T^*M, ω, H) . It follows that an equivalent way to define the geodesic flow is as follows.

Definition 5.24. Consider the smooth function

$$H : T^*M \rightarrow \mathbb{R} \quad (x, \xi) \mapsto \frac{1}{2} g^{ij}(x) \xi_i \xi_j$$

and its Hamiltonian vector field

$$V_H := \frac{\partial H}{\partial \xi_i} \frac{\partial}{\partial x^i} - \frac{\partial H}{\partial x^i} \frac{\partial}{\partial \xi_i}.$$

The flow generated by V_H is called the **geodesic flow**.

Example 5.25. (The Simple Pendulum Under the Legendre Transform)

Recall that in Example 3.25 we found the Euler-Lagrange equations for the simple pendulum. We also saw in Example 4.10 how it was described in the Hamiltonian formulation. Noticing that the pendulum was constrained to S^1 , we worked in a coordinate chart (U, θ) of S^1 and considered $(TS^1, \theta, \tilde{\theta})$. With the metric on S^1 induced from \mathbb{R}^2 , i.e. $g = g_{22} = l^2(d\theta)^2$, we saw that the kinetic energy was $K = \frac{1}{2} m l^2 (\tilde{\theta})^2$ and the potential energy was $U = mgl(1 - \cos \theta)$. Recall that in the Hamiltonian setting, we didn't change the potential, but we set the kinetic energy to be $K = \frac{\xi^2}{2ml^2}$. To see why we did this, we apply the Legendre transform to the Lagrangian set up. Let $(\Phi_L(TU), \theta, \xi)$ be a chart in the induced Hamiltonian system so that $\xi = \frac{\partial L}{\partial \tilde{\theta}} = \frac{\partial}{\partial \tilde{\theta}}(\frac{1}{2} m l^2 \tilde{\theta}^2) = m l^2 \tilde{\theta}$. It follows that $\tilde{\theta} = \frac{\xi}{m l^2}$. Thus,

$$K = \frac{1}{2} m l^2 \frac{\xi^2}{m^2 l^4} = \frac{\xi^2}{2 m l^2}.$$

Since in the induced Hamiltonian system the first coordinate is the same as in the Lagrangian setting, the potential energy remains unchanged.

6 Noether's Theorem

Noether's theorem provides a relationship between symmetries and constants of motion. Before going into the details we first need the formal definitions. In this section we will always assume that our manifolds are geodesically complete.

6.1 Noether's Theorem in Lagrangian Mechanics

First recall the different notions we have of 'lifting' maps. Let M be a manifold and $f : M \rightarrow M$ a diffeomorphism. In section 2.5 we defined the lift of f to the cotangent bundle to be the map

$$f_{\sharp} : T^*M \rightarrow T^*M \quad (x, \xi) \mapsto (f(x), (f^*)^{-1}(\xi)).$$

Note that we can also lift f to a map on the tangent bundle by taking the differential of f . To avoid confusion, we will denote the lift of f to TM by

$$\tilde{f} : TM \rightarrow TM \quad (p, V) \mapsto (f(p), f_{*,p}(V)).$$

Lastly, we defined in section 3.2 the lift of a curve $\gamma : \mathbb{R} \rightarrow M$ to the tangent bundle by

$$\tilde{\gamma} : \mathbb{R} \rightarrow TM \quad t \mapsto (\gamma(t), \gamma'(t)).$$

Definition 6.1. In a Lagrangian system (M, L) a **continuous symmetry** is a one parameter family of diffeomorphisms $\{\theta_s : M \rightarrow M ; s \in \mathbb{R}\}$ such that for each $s \in \mathbb{R}$ we have $(\tilde{\theta}_s)^*L = L$. That is, the family of maps is a continuous symmetry if for all $s \in \mathbb{R}$ we have that $L \circ \tilde{\theta}_s = L$.

A continuous symmetry can be thought of as a symmetry of motion. The standard notion of symmetry is invariance under some sort of mapping, i.e. an object is called symmetric if there is a map that preserves it. But here the objects being acted on are motions and so this definition is referring to the preservation of solutions to the Euler-Lagrange equations. Some obvious examples that we observe in homogeneous space are the invariance of the laws of motion under space and time translations. Noether's theorem says that both these families of continuous symmetries (space and time translations) have corresponding conserved quantities. We will see below that they are conservation of momentum and conservation of energy respectively.

Definition 6.2. In a Lagrangian system (M, L) a **conserved quantity** (or **constant of the motion**) is a smooth function $G \in C^\infty(TM)$ with the property that for any motion $\gamma : \mathbb{R} \rightarrow M$ in (M, L) , the total time derivative of G vanishes on the image of γ . That is for all $t \in \mathbb{R}$

$$\frac{d}{dt}G(\gamma(t)) = 0$$

All of the conservation laws in physics correspond to a conserved quantity.

Theorem 6.3. (Nöether) Let (M, L) be a Lagrangian system and $\gamma : \mathbb{R} \rightarrow M$ a motion. Let (x^1, \dots, x^n) be local coordinates for M and $(x^1, \dots, x^n, v^1, \dots, v^n)$ the induced local coordinates on TM . For any continuous symmetry $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ in (M, L) there exists a conserved quantity. The conserved quantity is given by the formula

$$\left(\frac{\partial}{\partial v^i} L \right) \cdot \left(\frac{d}{ds} \circ \theta_{s,*}^i \right) \in C^\infty(TM).$$

Proof. By hypothesis, for any $s \in \mathbb{R}$, $L(\widetilde{(\theta_s \circ \gamma)}(t)) = L(\widetilde{\gamma}(t))$. That is,

$$\begin{aligned} 0 &= \frac{\partial}{\partial s} L(\widetilde{(\theta_s \circ \gamma)}(t)) \\ &= \frac{\partial}{\partial s} L((\theta_s \circ \gamma)(t), (\theta_s \circ \gamma)'(t)) \\ &= \frac{\partial}{\partial x^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \left(\frac{d}{ds} (\theta_s \circ \gamma)^i(t) \right) + \frac{\partial}{\partial v^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \left(\frac{d}{ds} \left(\frac{d}{dt} (\theta_s \circ \gamma)^i(t) \right) \right) \end{aligned}$$

By hypothesis we have $\widetilde{\theta_s \circ \gamma}$ satisfies the Euler-Lagrange equations:

$$\frac{\partial}{\partial x^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) = \frac{d}{dt} \left(\frac{\partial}{\partial v^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \right)$$

Plugging in the left hand side of this equation into the above gives

$$\begin{aligned} 0 &= \frac{d}{dt} \left(\frac{\partial}{\partial v^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \right) \left(\frac{d}{ds} (\theta_s \circ \gamma)^i(t) \right) + \frac{\partial}{\partial v^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \left(\frac{d}{ds} \left(\frac{d}{dt} (\theta_s \circ \gamma)^i(t) \right) \right) \\ &= \frac{d}{dt} \left[\left(\frac{\partial}{\partial v^i} L(\widetilde{(\theta_s \circ \gamma)}(t)) \right) \left(\frac{d}{ds} (\theta_{s,*}(\widetilde{\gamma}(t)))^i \right) \right] \end{aligned}$$

That is, $\left(\frac{\partial}{\partial v^i} L \right) \cdot \left(\frac{d}{ds} \circ \theta_{s,*}^i \right)$ is a conserved quantity. \square

Remark 6.4. By noticing that $\frac{\partial L}{\partial v^i}$ is nothing but the Legendre transform of L , we can give a coordinate free description of the resulting conserved quantity. That is, setting

$$G := \left(\frac{\partial L}{\partial v^i} \right) \left(\frac{d}{ds} \circ (\theta_{s,*})^i \right)$$

we have that for arbitrary $(p, v) \in TM$

$$G(p, v) = (\Phi_{L_p}(v)) \cdot \left(\theta_*^{(p)} \left(\frac{d}{ds} \Big|_{s=0} \right) \right).$$

Example 6.5. (SO(3) gives Continuous Symmetries under a Central Force)

Let F be a central force acting on a particle in \mathbb{R}^3 with metric g and natural Lagrangian $L = K - U$. Let γ be the motion of the particle. A *one-parameter subgroup* of $SO(3)$ is a collection

of maps $\{\theta_s \in SO(3)\}$, where $s \in \mathbb{R}$, with the property that $\theta_{s+t} = \theta_s \circ \theta_t$. By definition, for every $x \in \mathbb{R}^3$ we have that $\|\theta_s(x)\| = \|x\|$ and hence for each $1 \leq i \leq 3$ we have that $U(\theta_s(\gamma(t))) = U(\gamma(t))$. Since each θ_s is a linear operator on \mathbb{R}^3 each map has a matrix representation, say $[\theta_s]$. Note that

$$\begin{aligned} v^i(\widetilde{(\theta_s \circ \gamma)}(t)) v^j(\widetilde{(\theta_s \circ \gamma)}(t)) &= \frac{d}{dt}((\theta_s \circ \gamma(t))^i) \cdot \frac{d}{dt}((\theta_s \circ \gamma(t))^j) \\ &= ([(\gamma^i)'(t)]^T [\theta_s]^T) \cdot ([\theta_s][(\gamma^i)'(t)]) && \text{since } \theta_s \text{ is linear} \\ &= v^i(\widetilde{\gamma}(t)) v^j(\widetilde{\gamma}(t)) && \text{since } [\theta_s] \text{ is orthogonal} \end{aligned}$$

It follows that $K(\widetilde{(\theta_s \circ \gamma)}(t)) = K(\widetilde{\gamma}(t))$. Hence $L(\widetilde{(\theta_s \circ \gamma)}(t)) = L(\gamma(t))$ showing that each one parameter subgroup of $SO(3)$ is a continuous symmetry on natural Lagrangian systems under a central force. This statement easily generalizes to \mathbb{R}^n . It also generalizes to a system of k particles in \mathbb{R}^n by considering the manifold \mathbb{R}^{nk} so that the motion of the k particles is described by one curve.

Remark 6.6. In a natural Lagrangian system for which the potential energy is zero, every element of the Gallilean group $\text{SGal}(3)$ corresponds to a continuous symmetry. In the cases where the Euler-Lagrange equations reduce to Newton's second law, saying that the elements of $\text{SGal}(3)$ are continuous symmetries is equivalent to the statement that Newton's laws are invariant under the action of elements of $\text{SGal}(3)$.

Example 6.7. (Rotational Invariance Gives Conservation of Angular Momentum)

Example 6.5 showed that, under a central force, each one parameter subgroup of $SO(3)$ is a continuous symmetry. Hence by Noether's theorem, each has a corresponding conserved quantity. For example consider rotation about the x^3 -axis. That is, consider the one parameter family given by

$$\theta_s = \begin{bmatrix} \cos s & -\sin s & 0 \\ \sin s & \cos s & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Noether's theorem shows that for any $t \in \mathbb{R}$ the corresponding conserved quantity is

$$\left(\Phi_{L_{\gamma(t)}}(\gamma'(t)) \right) \cdot \left(\theta_*^{(\gamma(t))} \left(\frac{d}{ds} \Big|_{s=0} \right) \right).$$

Computing, we get that the conserved quantity is

$$\begin{aligned} & [m(\gamma^1)'(t), m(\gamma^2)'(t), m(\gamma^3)'(t)] \cdot \left(\begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} (\gamma^1)'(t) \\ (\gamma^2)'(t) \\ (\gamma^3)'(t) \end{bmatrix} \right) \\ &= m(\gamma^2)'(\gamma^1)' - m(\gamma^1)'(\gamma^2)' \end{aligned}$$

This last value is precisely the angular momentum in the x^3 direction. That is, rotational symmetry about the x^3 -axis has angular momentum in the x^3 direction as its corresponding conserved quantity.

Example 6.8. (Translational Invariance Gives Conservation of Momentum)

Consider a closed system in $(\mathbb{R}^3, K - U)$ subject to a conservative force whose potential is independent of the x^1 -coordinate. Notice that translation in the x^1 -direction is given by the one parameter family $\{\theta_s : \mathbb{R} \rightarrow \mathbb{R}^3, (x^1, x^2, x^3) \mapsto (x^1 + s, x^2, x^3)\}$. For a motion γ , it's clear that $\frac{d}{dt}(\theta_s \circ \gamma(t)) = \gamma'(t)$ and so

$$\widetilde{(\theta_s \circ \gamma)}(t) = (\gamma^1(t) + s, \gamma^2(t), \gamma^3(t), (\gamma^1)'(t), (\gamma^2)'(t), (\gamma^3)'(t)).$$

By our hypothesis we have that $U(\theta_s(\gamma(t))) = U(\gamma(t))$. Hence $L(\widetilde{(\theta_s \circ \gamma)}(t)) = L(\widetilde{\gamma}(t))$. By Noether's theorem we have that

$$\left(\frac{\partial L}{\partial v^i}(\widetilde{(\theta_s \circ \gamma)}(t)) \right) \cdot \left(\frac{d}{ds} \Big|_{s=0} (\theta_{s,*}(\widetilde{\gamma}(t)))^i \right)$$

is a conserved quantity. But this is equal to

$$[m(\gamma^1)'(t), m(\gamma^2)'(t), m(\gamma^3)'(t)] \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = m(\gamma^1)'(t).$$

That is, momentum in the x^1 -direction is the resulting conserved quantity of translation in the x^1 -direction. This example easily generalizes to \mathbb{R}^n . Note also that in a system with k particles in \mathbb{R}^n interacting through conservative forces, we can replace our base space with \mathbb{R}^{nk} and study the motion of 1 particle.

Example 6.9. (Time Invariances Gives Conservation of Energy)

We showed in proposition 5.17 that if the Lagrangian was time independent then the total energy was conserved. However, it is not obvious how to view time translation as a continuous symmetry. To see how this can be done, suppose first that in a Lagrangian system (M, L) the Lagrangian is time dependent. That is, suppose $L \in C^\infty(T(M \times \mathbb{R}))$. Consider the Lagrangian system $(M \times \mathbb{R}, L_1)$ where $L_1 \in C^\infty(T(M \times \mathbb{R}))$ is defined as follows. If $(x^1, \dots, x^n, v^1, \dots, v^n)$ are the induced coordinates on TM and (t, u) are the induced coordinates on $T\mathbb{R}$ then define L_1 by

$$L_1 : T(M \times \mathbb{R}) \rightarrow \mathbb{R} \quad (t, x, u, v) \mapsto L(t, x, \frac{v}{u})u.$$

Notice that if we have a motion $\gamma : \mathbb{R} \rightarrow M$, then the curve

$$\hat{\gamma} := \alpha : \mathbb{R} \rightarrow \mathbb{R} \times M \quad \varepsilon \mapsto (\varepsilon, \gamma(\varepsilon))$$

is a motion in $(\mathbb{R} \times M, L_1)$. This is because $\tilde{\alpha}(\varepsilon) = (\varepsilon, \gamma(\varepsilon), 1, \gamma'(\varepsilon))$ and

$$L_1(\varepsilon, \gamma(\varepsilon), 1, \gamma'(\varepsilon)) = L(\varepsilon, \gamma(\varepsilon), \gamma'(\varepsilon)).$$

Hence it is justified to call $\gamma : \mathbb{R} \rightarrow M$ a motion in (M, L) if $\hat{\gamma} = \alpha$ is a motion in $(\mathbb{R} \times M, L_1)$. Suppose that $\{\theta_s : \mathbb{R} \times M \rightarrow \mathbb{R} \times M\}$ is a continuous symmetry in $(\mathbb{R} \times M, L_1)$. Then by Noether's theorem we have that the following is a conserved quantity in $(\mathbb{R} \times M, L_1)$;

$$\frac{\partial L_1}{\partial u}(t, x, u, v) \cdot \frac{d}{ds}\theta_s^0 + \frac{\partial L_1}{\partial v^i}(t, x, u, v) \cdot \frac{d}{ds}\theta_s^i$$

which is

$$L(t, x, \frac{v}{u}) \frac{d}{ds}\theta_s^0 - \frac{v^i}{u} \frac{\partial L}{\partial v^i}(t, x, \frac{v}{u}) \frac{d}{ds}\theta_s^0 + \frac{\partial L}{\partial v^i}(t, x, \frac{v}{u}) \frac{d}{ds}\theta_s^i.$$

If the Lagrangian is time dependent we define conserved quantities as follows. Given a motion γ in (M, L) we get the motion $\alpha = \hat{\gamma}$ in $(\mathbb{R} \times M, L_1)$ defined above. We observed that a continuous symmetry / conserved quantity in $(\mathbb{R} \times M, L_1)$ is also a continuous symmetries / conserved quantity in (M, L) since $L_1 = L$ when we set $u = 1$. By Noether's theorem, a continuous symmetry $\{\theta_s\}$ in $(\mathbb{R} \times M, L_1)$ gives a conserved quantity in $(\mathbb{R} \times M, L_1)$ which in turn gives a conserved quantity, G , in (M, L) by setting $u = 1$. That is we have that $\{\theta_s\}$ is a continuous symmetry in (M, L) with corresponding conserved quantity G . We can now view time translation as a continuous symmetry and compute its corresponding conserved quantity. It's clear that the continuous symmetry of $(\mathbb{R} \times M, L_1)$ given by $\{\theta_s : \mathbb{R} \times M \rightarrow \mathbb{R} \times M, (t, x, u, v) \mapsto (t + s, x, u, v)\}$ is representing time translation. Now suppose that $\gamma : \mathbb{R} \rightarrow M$ is a motion in (M, L) and L is time independent. Consider the induced motion $\hat{\gamma} = \alpha : \mathbb{R} \rightarrow \mathbb{R} \times M$. We have that

$$\begin{aligned} L_1(\widetilde{(\theta_s \circ \alpha)})(t) &= L_1(\varepsilon + s, \gamma(\varepsilon), 1, \gamma'(\varepsilon)) \\ &= L(\varepsilon + s, \gamma(\varepsilon), \gamma'(\varepsilon)) \\ &= L(\varepsilon, \gamma(\varepsilon), \gamma'(\varepsilon)) && \text{since } L \text{ is time independent} \\ &= L_1(\varepsilon, \gamma(\varepsilon), 1, \gamma'(\varepsilon)) \\ &= L_1(\tilde{\alpha}) \end{aligned}$$

But we have that $\frac{d}{ds}\theta_s^i = 0$ for all $i \geq 1$. As well, $\frac{d}{ds}\theta_s^0 = 1$ and so the conserved quantity is just $L(t, x, \frac{v}{u}) - \frac{v^i}{u} \frac{\partial L}{\partial v^i}(t, x, \frac{v}{u})$. By setting $u = 1$, it follows that the conserved quantity in (M, L) is $L(t, x, v) - v^i \frac{\partial L}{\partial v^i}(t, x, v) = -L^*$. But in a natural Lagrangian system, Proposition 5.15 showed that L^* is the total energy.

6.2 Noether's Theorem in Hamiltonian Mechanics

Fix a Hamiltonian system (X, ω, H) . We first define the notions of continuous symmetry and conserved quantity in the symplectic setting.

Definition 6.10. A **continuous symmetry** is a vector field $W \in \Gamma(TX)$ such that

$$\mathcal{L}_W \omega = 0 \quad \text{and} \quad \mathcal{L}_W H = 0$$

Definition 6.11. A **conserved quantity** (or **constant of motion**) is a function $f \in C^\infty(X)$ that Poisson commutes with H . By the antisymmetry of the Poisson bracket, this means that

$$\{f, H\} = 0 = \{H, f\}.$$

Since $\{H, f\} = \omega(V_H, V_f) = (V_H \lrcorner \omega)(V_f) = (dH)(V_f) = V_f H = \mathcal{L}_{V_f} H$, we have that f Poisson commutes with H if and only if

$$\mathcal{L}_{V_f} H = V_f H = 0 = V_H f = \mathcal{L}_{V_H} f.$$

Theorem 6.12. (Noether) If $W \in \Gamma(TX)$ is a continuous symmetry, then W is locally Hamiltonian and its Hamiltonian function is a constant of the motion. Conversely, given a constant of the motion $f \in C^\infty(X)$, its Hamiltonian vector field, V_f , is a continuous symmetry.

Proof. Let W be a continuous symmetry. By hypothesis,

$$0 = \mathcal{L}_W \omega = d(W \lrcorner \omega) + W \lrcorner d\omega = d(W \lrcorner \omega)$$

Hence, by Poincaré's lemma, around every point there exists a neighbourhood U and a function $f \in C^\infty(U)$ such that $W \lrcorner \omega = df$. That is, locally $W = V_f$ so that W is locally Hamiltonian. By hypothesis, $\mathcal{L}_W H = 0 = \mathcal{L}_{V_f} H$ so that $\{f, H\} = 0 = \{H, f\}$. Conversely, let $f \in C^\infty(X)$ be a conserved quantity so that $\{f, H\} = 0 = \{H, f\}$. Consider the corresponding vector field V_f . It was just shown that $\mathcal{L}_{V_f} H = 0$ while

$$\mathcal{L}_{V_f} \omega = d(V_f \lrcorner \omega) + V_f \lrcorner d\omega = d(V_f \lrcorner \omega) = d(df) = 0$$

□

Example 6.13. (Symmetries on the 2-Torus)

Consider the 2-torus \mathbb{T}^2 with local coordinate chart (U, θ, φ) . Here $U = A \times B$, where (A, θ) and (B, φ) are two different local coordinate charts on S^1 . Consider the Hamiltonian system $(U, \omega = d\theta \wedge d\varphi, H)$ where $H \in C^\infty(TU)$ is defined by $H(\theta, \varphi) := \theta$. It follows $dH = d\theta$ and so the Hamiltonian vector field is $V_H = -\frac{\partial}{\partial \varphi}$. Consider the vector field $W = \frac{\partial}{\partial \varphi} \in \Gamma(TU)$. The flow of W is

$$\theta_s : U \rightarrow U, \quad (\theta, \varphi) \mapsto (\theta, \varphi + s).$$

By Cartan's magic formula, we have that

$$\mathcal{L}_{\frac{\partial}{\partial \varphi}} \omega = d(d\theta) + \frac{\partial}{\partial \varphi} \lrcorner d\omega = 0$$

and

$$\mathcal{L}_W H = \frac{\partial \theta}{\partial \varphi} = 0.$$

That is, W is a continuous symmetry in this Hamiltonian system. Since $W \lrcorner \omega = d\theta$ it follows that on U we have $W = V_f$ where $f : U \rightarrow \mathbb{R}$, $(\theta, \varphi) \mapsto \theta$. That is, $f = H$ and hence $\{f, H\} = 0$. That is, f is the corresponding conserved quantity. Notice that $f = H$ is not a global function on \mathbb{T}^2 and so we can only say that W is locally Hamiltonian.

Noether's theorem in the Hamiltonian setting says that each continuous symmetry on a symplectic manifold is locally Hamiltonian. That is, if $W \in \Gamma(TX)$ is a continuous symmetry, then for each $p \in X$ there exists an open set U_p and $f_p \in C^\infty(U_p)$ (the conserved quantity) such that $W = V_{f_p}$ on U_p . However, when the symplectic manifold is a cotangent bundle, the following proposition gives a condition on when the conserved quantity is global.

Proposition 6.14. *Given a manifold M , we know that the cotangent bundle $(T^*M, \omega = -d\alpha)$ is a symplectic manifold, where ω is the tautological 2-form. Fix a Hamiltonian function $H \in C^\infty(T^*M)$. Let $W \in \Gamma(T(T^*M))$ be a continuous symmetry in the Hamiltonian system (T^*M, ω, H) which preserves that tautological 1-form α . Then the corresponding conserved quantity (Hamiltonian function) is $\alpha(W)$. Note that this conserved quantity is globally defined. In other words, if $W \in \Gamma(T(T^*M))$ is a continuous symmetry which preserves α , then*

$$W = V_{\alpha(W)}.$$

Proof. Suppose that $W \in \Gamma(T(T^*M))$ is a continuous symmetry such that $\mathcal{L}_W \alpha = 0$. By Cartan's magic formula this means that

$$d(W \lrcorner \alpha) = -W \lrcorner d\alpha = W \lrcorner \omega.$$

□

6.3 Noether's Theorem Under the Legendre Transform

This subsection gives another example of how the Legendre transform translates statements between Lagrangian and Hamiltonian mechanics. That is, we show how the Legendre transform translates the statements of Noether's theorem.

Theorem 6.15. *Let (M, L) be a Lagrangian system where $L \in C^\infty(TM)$ is strongly convex. Suppose that $\{\theta_s : M \rightarrow \mathbb{R}, s \in \mathbb{R}\}$ is a continuous symmetry so that the corresponding conserved quantity is $F := \left(\frac{\partial}{\partial v^i} L\right) \left(\frac{d}{ds} \circ (\theta_{s,*}^i)\right)$. This continuous symmetry generates a vector field $W \in \Gamma(TM)$. The claim is that the vector field $W_\sharp \in \Gamma(T(T^*M))$, as defined in Lemma 2.24, is a continuous symmetry in the induced Hamiltonian system $(T^*M, \omega, H = L^*)$ and the corresponding conserved quantity is $F \circ \Phi_L^{-1} = F \circ \Phi_H$.*

Conversely, given an arbitrary manifold M , consider the Hamiltonian system $(T^*M, \omega = -d\alpha, H)$ for some arbitrary strongly convex Hamiltonian $H \in C^\infty(T^*M)$. Suppose that $W \in \Gamma(T(T^*M))$ is a continuous symmetry which preserves α . Let G denote the corresponding conserved quantity. If $\{\psi_s : T^*M \rightarrow T^*M, s \in \mathbb{R}\}$ is the flow of W then by Theorem 2.25 there exists a family of diffeomorphisms $\{\theta_s : M \rightarrow M\}$ such that $\theta_s^\sharp = \psi_s$. The claim is that $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ is a continuous symmetry in $(M, L := H^*)$ and that the corresponding conserved quantity is $G \circ \Phi_H^{-1} = G \circ \Phi_L$.

Proof. Given a continuous symmetry $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ in (M, L) , let $F = \left(\frac{\partial L}{\partial v^i}\right) \left(\frac{d}{ds} \circ \theta_{s,*}^i\right)$ be the corresponding conserved quantity. Let W be the infinitesimal generator of $\{\theta_s\}$ and W_\sharp denote its lift. We are trying to show that W_\sharp is a continuous symmetry in (T^*M, ω, H) . By Proposition 2.22 it follows

$$\mathcal{L}_{W_\sharp} \alpha = 0 \quad \text{and} \quad \mathcal{L}_{W_\sharp} \omega = 0$$

It remains to show that $\mathcal{L}_{W_\sharp} H = 0$. By Theorem 6.14 we know that $W_\sharp = V_{\alpha(W_\sharp)}$. However, $W_\sharp \lrcorner \alpha = F \circ \Phi_L^{-1} = F \circ \Phi_H$ since for arbitrary $(p, \xi_p) \in T^*M$ by definition

$$\begin{aligned} F \circ \Phi_L^{-1}(p, \xi_p) &= (\Phi_L(\Phi_L^{-1}(\xi_p))) \left(\theta_*^{(p)} \left(\frac{d}{ds} \Big|_{s=0} \right) \right) \\ &= \xi_p(W_p) \\ &= \xi_p(\pi_*((W_\sharp)_{(p, \xi_p)})) \\ &= \alpha(W_\sharp)(p, \xi_p) \end{aligned}$$

Thus, showing that $\mathcal{L}_{W_\sharp} H = 0$ is equivalent to showing, by Proposition 4.12, that for any integral curve Ψ of V_H we have $\frac{d}{dt}(F \circ \Phi_L^{-1})(\Psi(t)) = 0$. But by Theorem 5.17, any integral curve of V_H is of the form $(\Phi_L \circ \tilde{\alpha})(t)$ for some motion $\alpha(t)$ in (M, L) . But then

$$\frac{d}{dt} F \circ \Phi_L^{-1}(\Phi_L(\alpha(t))) = \frac{d}{dt} F(\alpha(t)) = 0$$

since F is a conserved quantity. All of this shows that W_\sharp is a conserved quantity in (T^*M, ω, L^*) with globally defined conserved quantity $F \circ \Phi_H$.

Conversely, consider a Hamiltonian system of the form $(T^*M, \omega = -d\alpha, H)$ for some strongly convex $H \in C^\infty(T^*M)$. Let $W \in \Gamma(T(T^*M))$ be a continuous symmetry whose flow preserves α . Then by Theorem 6.14 we have that the corresponding conserved quantity is $G = W \lrcorner \alpha$. Furthermore, by Theorem 2.25, if $\{\Psi_s : T^*M \rightarrow T^*M, s \in \mathbb{R}\}$ is the flow of W then there exists a family $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ such that $\Psi_s = \theta_s^\sharp$. That is, each Ψ_s is defined by

$$\Psi_s : T^*M \rightarrow T^*M \quad (p, \xi) \mapsto (\theta_s(p), (\theta_s^*)^{-1}(\xi)).$$

We want to show that the family $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ is a continuous symmetry in the induced

Lagrangian setting $(M, L = H^*)$. Consider the vector field

$$Y = \xi_i \frac{\partial}{\partial \xi_i} \in \Gamma(T(T^*M)),$$

which is negative the symplectic dual of α . By Claim 2.26 we have that

$$\Psi_{s,*}Y = Y,$$

while by hypothesis

$$\Psi_s^*H = H.$$

By definition

$$\begin{aligned} L(x, v) &= H^*(x, v) \\ &= \xi_i \frac{\partial H}{\partial \xi_i} - H(x, \xi) \\ &= Y(H) - H(x, \xi). \end{aligned}$$

Hence

$$\begin{aligned} (\Psi_s^*L)(x, v) &= \Psi_s^*(Y(H)) - \Psi_s^*(H(x, \xi)) \\ &= \Psi_s^*(Y(H)) - H(x, \xi) \end{aligned}$$

Thus it suffices to show that $\Psi_s^*(Y(H)) = Y(H)$. Indeed, for arbitrary $p \in M$ we have that

$$\begin{aligned} (\Psi_s^*(Y(H)))_p &= ((Y(H)) \circ \Psi_s)_p \\ &= Y_{\Psi_s(p)}H \\ &= (dH)_{\Psi_s(p)}(Y_{\Psi_s(p)}) \\ &= (dH)_{\Psi_s(p)}((\Psi_{s,*}Y)_p) && \text{since } \Psi_{s,*}Y = Y \\ &= (\Psi_s^*(dH))_p(Y_p) \\ &= (d(\Psi_s^*H))_p(Y_p) \\ &= (dH)_p(Y_p) && \text{since } \Psi_s^*H = H \\ &= Y_p(H) \end{aligned}$$

But by hypothesis, Ψ_s is the inverse of the pullback of θ_s . Thus we have shown that $(\tilde{\theta}_s)^*L = L$ for all $s \in \mathbb{R}$. That is, $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ is a continuous symmetry in the Lagrangian system (M, L) . Noether's theorem shows that the corresponding conserved quantity is $F = \left(\frac{\partial L}{\partial v^i}\right) \left(\frac{d}{ds} \circ \theta_{s,*}^i\right)$. As in the proof of the converse, we have that $F \circ \Phi_H = \alpha(X)$. That is

$$F = \alpha(X) \circ \Phi_L.$$

□

Remark 6.16. In summary we have shown that, given a continuous symmetry in (M, L) with corresponding conserved quantity F , there is a corresponding continuous symmetry in the Hamiltonian system $(T^*M, \omega, H = L^*)$ which preserves α and has conserved quantity $F \circ \Phi_H$. Conversely, given a continuous symmetry in a Hamiltonian system of the form (T^*M, ω, H) which preserves α and has corresponding conserved quantity G , there is a corresponding continuous symmetry in $(M, L = H^*)$ with conserved quantity $G \circ \Phi_L$. In subsection 6.5 we will consider what happens when we relax the definitions of continuous symmetry and conserved quantities.

Example 6.17. (Translational Invariance and Conservation of Momentum)

In example 6.8 we considered the natural Lagrangian system $(\mathbb{R}^3, K - U)$, where U was assumed to be independent of x^1 , with the continuous symmetry

$$\{\theta_s : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x^1, x^2, x^3) \mapsto (x^1 + s, x^2, x^3)\}.$$

We saw that as a consequence of this symmetry we got conservation of momentum. We can also see this by using Theorem 6.15, and converting to the Hamiltonian setting. Indeed our continuous symmetry generates the vector field $\frac{\partial}{\partial x^1} \in \Gamma(T\mathbb{R}^3)$. Since the Jacobian of θ_s is the identity matrix we have that $(\frac{\partial}{\partial x^1})_\#$ is just $\frac{\partial}{\partial x^1} \in \Gamma(T(T^*\mathbb{R}^3))$. By Theorem 6.15, it follows that $(\frac{\partial}{\partial x^1})_\#$ is a continuous symmetry with conserved quantity $\alpha\left((\frac{\partial}{\partial x^1})_\#\right) = \xi_1$. But ξ_1 is just mv^1 , which when applied to a motion is the momentum in the x^1 -direction.

6.4 The Converse of Noether's Theorem in the Lagrangian Setting

In subsection 6.1 we showed that given a Lagrangian system (M, L) and a continuous symmetry $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$, Noether's theorem gave the corresponding conserved quantity

$$\left(\frac{\partial L}{\partial v^i}\right) \cdot \left(\frac{d}{ds} \circ \theta_{s,*}^i\right).$$

With the results from the previous subsection we can now show that given a conserved quantity of this form, the corresponding family $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ is a continuous symmetry in (M, L) . Indeed, suppose that we have a family of diffeomorphisms $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ such that

$$\frac{d}{dt} \left(\left(\frac{\partial L}{\partial v^i} \right) \cdot \left(\frac{d}{ds} \circ \theta_{s,*}^i \right) \right) = 0.$$

This family of diffeomorphisms generates a vector field $Y \in \Gamma(TM)$. In Theorem 6.15, we showed that the vector field $Y_\# \in \Gamma(T(T^*M))$ is a continuous symmetry in $(T^*M, \omega, H = L^*)$ which preserves α and has conserved quantity $\alpha(Y_\#)$. However, by Lemma 2.24 the flow of $Y_\#$ is $\theta_{s,\#}$. It follows from the proof of Theorem 6.15 that $\{\theta_s : M \rightarrow M, s \in \mathbb{R}\}$ is a continuous symmetry

in $(M, H^*) = (M, L)$. Hence, we have the following diagram, where within each brace we are considering the induced Hamiltonian or Lagrangian system:

$$\begin{array}{ccc}
\left\{ \begin{array}{l} \text{Families of diffeomorphisms} \\ \{\theta_s : M \rightarrow M, s \in \mathbb{R}\} \\ \text{such that } (\tilde{\theta}_s)^* L = L \end{array} \right\} & \longleftrightarrow & \left\{ \begin{array}{l} X \in \Gamma(T(T^*M)) \\ \text{such that} \\ 0 = \mathcal{L}_X \omega = \mathcal{L}_X \alpha = \mathcal{L}_X H \end{array} \right\} \\
\downarrow & & \downarrow \\
\left\{ \begin{array}{l} \text{Conserved quantities} \\ \text{of the form } \left(\frac{\partial L}{\partial v^i} \right) \left(\frac{d}{ds} \circ \theta_{s,*}^i \right) \\ \text{for some family of maps} \\ \{\theta_s : M \rightarrow M\} \end{array} \right\} & \longleftrightarrow & \left\{ \begin{array}{l} f \in C^\infty(T^*M) \text{ with } f = \alpha(W) \\ \text{for some } W \in \Gamma(T(T^*M)) \\ \text{such that } \mathcal{L}_W \alpha = 0 \text{ and } \{f, H\} = 0 \end{array} \right\}
\end{array}$$

This diagram proves the converse of Noether's theorem in the Lagrangian setting when the continuous symmetries are restricted to be of the above form. Using this diagram we can exhibit a conserved quantity in a Lagrangian system which does not arise via Noether's theorem from a continuous symmetry.

Example 6.18. (The Laplace-Runge-Lenz Vector)

Consider a particle of mass m moving under a central force field in \mathbb{R}^3 . Let \vec{r} denote the position vector of this particle. For simplicity, endow \mathbb{R}^3 with the standard metric and let (x^1, x^2, x^3) denote the standard coordinates on \mathbb{R}^3 and $(x^1, x^2, x^3, v^1, v^2, v^3)$ and $(x^1, x^2, x^3, \xi_1, \xi_2, \xi_3)$ the induced coordinates on $T\mathbb{R}^3$ and $T^*\mathbb{R}^3$ respectively. Let $L = K - U$ be the natural Lagrangian. As is the case for the gravitational and electrostatic forces, we assume our potential energy is of the form $U = -\frac{k}{r}$ where k is some constant. As in the proof of Proposition 5.14 we have that locally $\xi_i = mv^i$, the momentum in the i th direction of the particle. That is, $\xi_i = p_i$ as functions on $\mathbb{R}^6 = T^*\mathbb{R}^3 = T\mathbb{R}^3$. In these coordinates we have that the tautological one form is $\alpha = p_i dx^i$. The **Laplace-Runge-Lenz** vector is defined to be

$$\vec{A} = \vec{p} \times \vec{L} - \frac{mk\vec{r}}{\|\vec{r}\|}$$

where \vec{p} is the particle's momentum and \vec{L} is the particle's angular momentum. By Proposition 3.11 we have that the $\frac{d}{dt}\vec{L} = 0$. Identifying the vector field $\frac{\partial}{\partial r}$ with $\frac{\vec{r}}{\|\vec{r}\|}$ we have that

$$ma = m\ddot{\vec{r}} = F = -\nabla U = -\frac{\partial U}{\partial r} \frac{\partial}{\partial r} = -\frac{k}{\|r\|^3} \vec{r}.$$

It follows that

$$\begin{aligned}
\frac{d\vec{A}}{dt} &= \dot{\vec{p}} \times \vec{L} + \vec{p} \times \dot{\vec{L}} - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= \dot{\vec{p}} \times \vec{L} - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= -\frac{mk}{\|\vec{r}\|^3} \left(\vec{r} \times (\vec{r} \times \dot{\vec{r}}) \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= -\frac{mk}{\|\vec{r}\|^3} \left((\vec{r} \cdot \dot{\vec{r}}) \vec{r} - \|\vec{r}\|^2 \dot{\vec{r}} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= -\frac{mk}{\|\vec{r}\|^3} \left(\frac{1}{2} \frac{d}{dt} (\vec{r} \cdot \vec{r}) \vec{r} - \|\vec{r}\|^2 \dot{\vec{r}} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= -\frac{mk}{\|\vec{r}\|^3} \left(\frac{1}{2} \frac{d}{dt} (\|\vec{r}\|^2) \vec{r} - \|\vec{r}\|^2 \dot{\vec{r}} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= -\frac{mk}{\|\vec{r}\|^3} \left(\|\vec{r}\| \left(\frac{d}{dt} \|\vec{r}\| \right) \vec{r} - \|\vec{r}\|^2 \dot{\vec{r}} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= mk \left(\frac{\dot{\vec{r}}}{\|\vec{r}\|} - \frac{\left(\frac{d}{dt} \|\vec{r}\| \right)}{\|\vec{r}\|^2} \vec{r} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) - \frac{d}{dt} \left(\frac{mk}{\|\vec{r}\|} \vec{r} \right) \\
&= 0
\end{aligned}$$

Hence the first component of \vec{A} , which is

$$A^1 = (\vec{p} \times \vec{L})^1 - \frac{mkx^1}{\|\vec{r}\|} = p^2 L^3 - p^3 L^2 - \frac{mkx^1}{\|\vec{r}\|}$$

is a conserved quantity. That is, $A^1 \in C^\infty(\mathbb{R}^6)$ is such that $\frac{d}{dt} A^1 = 0$ on motions in the Hamiltonian system. It follows, from Theorem 6.15 that $A^1 \circ \Phi_L$ is a conserved quantity in the Lagrangian system $(\mathbb{R}^3, K - U)$. In order to show that this conserved quantity does not arise from a continuous symmetry, by the above diagram we need to show that the induced continuous symmetry in the Hamiltonian system $(T^*\mathbb{R}^3, K + U)$ corresponding to the conserved quantity $A^1 \circ \Phi_L \circ \Phi_H = A^1$ has flow which is not the lift of curves on the base manifold.

By definition, our Hamiltonian system is $(T^*\mathbb{R}^3, dx^i \wedge dp_i, L^*)$, where by the argument in the proof of Proposition 5.15 we have that

$$K + U = H := L^* = \frac{\|p\|^2}{2m} - \frac{k}{\|r\|}.$$

Note that by definition

$$\begin{aligned}
X_{A^1} &= \frac{\partial A^1}{\partial p_i} \frac{\partial}{\partial x^i} - \frac{\partial A^1}{\partial x^i} \frac{\partial}{\partial p_i} \\
&= L^3 \frac{\partial}{\partial x^2} - L^2 \frac{\partial}{\partial x^3} - \frac{mk}{r} \frac{\partial}{\partial p_1} + \frac{mkx^1}{\|\vec{r}\|^2} \frac{\partial r}{\partial x^i} \frac{\partial}{\partial p_i} \\
&= L^3 \frac{\partial}{\partial x^2} - L^2 \frac{\partial}{\partial x^3} + \frac{mk}{\|\vec{r}\|^3} (-(x^2)^2 - (x^3)^2) \frac{\partial}{\partial p_1} + \frac{mk}{\|\vec{r}\|^3} x^1 x^2 \frac{\partial}{\partial p_2} + \frac{mk}{\|\vec{r}\|^3} x^1 x^3 \frac{\partial}{\partial p_3}
\end{aligned}$$

A straightforward calculation shows that $X_{A^1}H = 0$ so that A^1 is a conserved quantity in the induced Hamiltonian setting.

To find the flow, or integral curves $\gamma(t) = (x(t), \xi(t))$, of the continuous symmetry X_{A^1} we need to solve the Hamilton equations:

$$\begin{aligned}
(x^1)' &= 0 \\
(x^2)' &= L^3 \\
(x^3)' &= -L^2 \\
(p_1)' &= -\frac{mk}{\|\vec{r}\|^3} (-(x^2)^2 - (x^3)^2) \\
(p_2)' &= \frac{mk}{\|\vec{r}\|^3} x^1 x^2 \\
(p_3)' &= \frac{mk}{\|\vec{r}\|^3} x^1 x^3
\end{aligned}$$

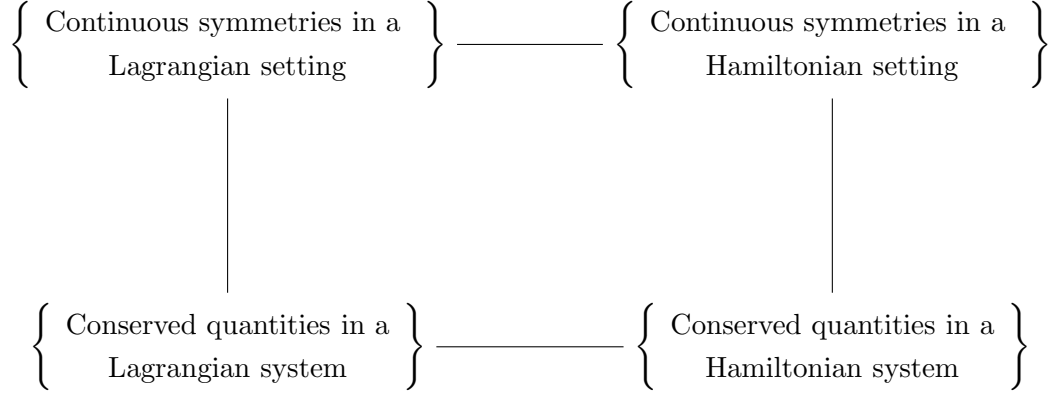
All we are trying to show is that the flow of X_{A^1} is not the lift of a one parameter family of diffeomorphisms on M . But by definition, given a continuous symmetry $\theta_s : M \rightarrow M$, the lift $\theta_{s,\sharp}$ is equal to $(\theta_s^*)^{-1}$, which is a linear function on each fibre of the cotangent bundle. In particular then, if the flow of X_{A^1} came from lifting curves on M it would be that each $(p_k)'$, for $k = 1, 2, 3$, were a linear function of p_1, p_2 and p_3 . However, we can see immediately from the form of the above ODE's that this is not the case. Thus the flow of X_{A^1} is not the lift of a continuous symmetry on \mathbb{R}^3 . We thus have shown that $A^1 \circ \Phi_L$ is a conserved quantity in the Lagrangian system $(\mathbb{R}^3, K - U)$ which does not come from a continuous symmetry.

Remark 6.19. An equivalent way to see this would be to show that $\mathcal{L}_{X_{A^1}}\alpha \neq 0$. This can be done explicitly, but the calculation is quite lengthy. Note that although $\mathcal{L}_{X_{A^1}}\alpha \neq 0$, we showed that X_{A^1} is a continuous symmetry in $(T^*\mathbb{R}^3, \omega, K + U)$ so that $\mathcal{L}_{X_{A^1}}\omega = 0$. That is, not all continuous symmetries in Hamiltonian systems of the form (T^*M, ω, H) need to preserve the tautological 1-form.

6.5 Relaxing the Definitions of Symmetries and Conserved Quantities

With the Laplace-Runge-Lenz vector in mind, it is an interesting question to consider what would happen if we relaxed the definitions of continuous symmetry and conserved quantity. In particular,

is there a way to make the following diagram traceable both clockwise and counterclockwise?



We can see right away that with our definitions this is impossible. Indeed, any continuous symmetry in a Lagrangian system (M, L) gives a continuous symmetry in (T^*M, ω, L) which preserves α . However, we showed above that the Laplace-Runge-Lenz vector is a continuous symmetry which does not preserve α .

Recall that the original definition of a continuous symmetry in the Hamiltonian setting does not include the requirement of preserving the tautological 1-form. However, even without this requirement we saw that, in the Hamiltonian setting, continuous symmetries and locally defined conserved quantities are in one-to-one correspondence. If we return to the Lagrangian setting with the original definitions of symmetries and conserved quantity, we see that the two notions are not in one-to-one correspondence. Indeed, we showed in the previous subsection that the Laplace-Runge-Lenz vector is a conserved quantity, not of the form $(\frac{\partial L}{\partial v^i})(\frac{d}{ds}\theta_{s,*}^i)$, which does not come from a continuous symmetry.

However, notice that the one-to-one correspondence of conserved quantities in the Lagrangian and Hamiltonian setting still holds. This is because if $F \in C^\infty(TM)$ is constant on the motions in (M, L) then we have, by Theorem 5.19, that

$$\frac{d}{dt}(F \circ \Phi_H)(\Psi(t)) = 0$$

for all motions $\Psi : \mathbb{R} \rightarrow T^*M$. That is, by Proposition 4.12, $F \circ \Phi_H$ is a conserved quantity in $(T^*M, \omega, H = L^*)$. Conversely, if $G \in C^\infty(T^*M)$ is such that $\{G, H\} = 0$ then $\frac{d}{dt}(G(\Psi(t))) = 0$ for all integral curves Ψ of X_H , then by Theorem 5.20 it follows that

$$\frac{d}{dt}(G \circ \Phi_L(\tilde{\gamma}(t))) = 0$$

for all motions $\gamma : \mathbb{R} \rightarrow M$ in (M, L) . That is we have the one-to-one correspondence

$$\left\{ \begin{array}{c} \text{Conserved quantities in a} \\ \text{Lagrangian setting} \end{array} \right\} \longleftrightarrow \left\{ \begin{array}{c} \text{Conserved quantities in a} \\ \text{Hamiltonian setting} \end{array} \right\}.$$

However, in order to reconcile the one-to-one correspondence between symmetries one needs to study the notion of ‘generalized symmetries’. A thorough treatment of this topic can be found in chapter 5 of [9]. Roughly speaking, these are a one-parameter family of ‘Lagrangian preserving’ maps in a Lagrangian system (M, L) which do not necessarily arise from lifting curves on M . That is, a generalized symmetry can be thought of as a one-parameter family $\{\Upsilon_s : TM \rightarrow TM, s \in \mathbb{R}\}$ such that $\Upsilon_s^* L = L$. Notice that with this new definition, Noether’s theorem still holds as stated in the Lagrangian setting. The proof of Theorem 6.3 did not use the fact that we were lifting each θ_s to the tangent bundle. If one replaces $\tilde{\theta}_s$ with Υ_s , the proof of Theorem 6.3 is unchanged. It is likely that with the notion of generalized symmetries, one can show that all arrows in the following diagram go both ways. We have filled in the correspondences discussed in this paper.

$$\begin{array}{ccc} \left\{ \begin{array}{c} \text{Families of diffeomorphisms} \\ \{\Upsilon_s : TM \rightarrow TM, s \in \mathbb{R}\} \\ \text{such that } (\Upsilon_s)^* L = L \end{array} \right\} & \longleftrightarrow & \left\{ \begin{array}{c} X \in \Gamma(T(T^*M)) \\ \text{such that} \\ \mathcal{L}_X \omega = 0 = \mathcal{L}_X H \end{array} \right\} \\ \downarrow & & \uparrow \\ \left\{ \begin{array}{c} \text{Conserved quantities} \\ F \in C^\infty(TM) \\ \text{such that } \frac{d}{dt} F = 0 \end{array} \right\} & \longleftrightarrow & \left\{ \begin{array}{c} f \in C^\infty(T^*M) \\ \text{such that} \\ \text{locally } \{f, H\} = 0 \end{array} \right\} \end{array}$$

References

- [1] Ana Cannas da Silva. *Lectures on Symplectic Geometry*, Springer-Verlag, 2001.
- [2] V.I. Arnold. *Mathematical Methods of Classical Mechanics*, 2nd edition, Springer-Verlag, 1989
- [3] Herbert Goldstein. *Classical Mechanics*, 2nd edition, Addison-Wesley Publishing Company, 1981.
- [4] John M. Lee. *Introduction to Smooth Manifolds*, 2nd edition, Springer, 2000.
- [5] Peter J. Olver. *Applications of Lie Groups to Differential Equations*, 2nd edition, Springer-Verlag, 1986.
- [6] John R. Taylor. *Classical Mechanics*, University Science Books, 2005.
- [7] Loring W. Tu. *An Introduction to Manifolds*, 2nd edition, Springer, 2011.
- [8] Ralph Abraham. *Foundations of Mechanics*, W.A. Benjamin, 1967.
- [9] Peter J. Olver. *Applications of Lie Groups to Differential Equations*, 2nd edition, Springer-Verlag, 1986.
- [10] John M. Lee. *Riemannian Manifolds: An Introduction to Curvature*, Springer-Verlag, 1997.
- [11] David Tong. *Lectures on Classical Dynamics*, <http://www.damtp.cam.ac.uk/user/tong/dynamics.htm>, University of Cambridge, 2012.
- [12] Jeremy Butterfield. *On Symmetry and Conserved Quantities in Classical Mechanics*, University of Western Ontario Series in Philosophy of Science, 2006
- [13] Jordan Bell. *The Legendre Transform*, <http://individual.utoronto.ca/jordanbell/notes/legendre.pdf>, University of Toronto, 2014
- [14] Stephen H. Friedberg, Arnold J. Insel, Lawrence E. Spence. *Linear Algebra*, 4th edition, Prentice Hall, 2003.